

# MIMO Gaussian Broadcast Channels with Common, Private and Confidential Messages

Ziv Goldfeld and Haim H. Permuter

## Abstract

The two-user multiple-input multiple-output (MIMO) Gaussian broadcast channel (BC) with common, private and confidential messages is considered. The transmitter sends a common message to both users, a confidential message to User 1 and a private (non-confidential) message to User 2. The secrecy-capacity region is characterized by showing that certain inner and outer bounds coincide and that the boundary points are achieved by Gaussian inputs. The proof relies on factorization of upper concave envelopes and a variant of dirty-paper coding (DPC). It is shown that the entire region is exhausted by using DPC to cancel out the signal of the non-confidential message at Receiver 1, thus making DPC against the signal of the confidential message unnecessary. A numerical example visualizes the secrecy-capacity results.

## Index Terms

Additive Gaussian channel, broadcast channel, dirty-paper coding multiple-input multiple-output (MIMO) communications, physical-layer security, upper concave envelopes.

## I. INTRODUCTION

Additive Gaussian channels are a common model for wireless communication, whose open nature makes it vulnerable to a variety of security threats, such as eavesdropping. However, eavesdroppers are not always a malicious entity from which *all* transmissions are concealed. Rather, a legitimate recipient of one message may serve as an eavesdropper for other messages. We encapsulate this notion in a two-user multiple-input multiple-output (MIMO) Gaussian broadcast channel (BC) with common, private and confidential messages (Fig. 1). The common message  $M_0$  is intended to both users, while  $M_1$  and  $M_2$  are private messages that are sent to users 1 and 2, respectively. Furthermore,  $M_1$  is confidential and is kept secret from user 2. Many real-life scenarios adhere to this setup. One such example is a banking site that simultaneously: (i) broadcasts an advertisement to all online users (modeled by  $M_0$ ); (ii) offers public information (such as material on different banking programs, reports, forecasts, etc.) that is available only to users that are interested in it (modeled by the private message  $M_2$ ); (iii) provides an online banking service, where users can access their account and perform transactions (this confidential information is modeled by  $M_1$ ).

The work of Z. Goldfeld and H. H. Permuter was supported by the Israel Science Foundation (grant no. 2012/14), the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n°337752, and the Cyber Center and at Ben-Gurion University of the Negev.

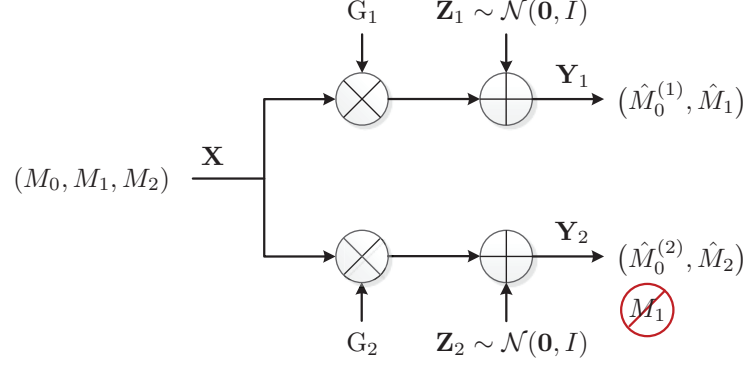


Fig. 1: MIMO Gaussian BC.

In recent years, information-theoretic security over MIMO communication systems has been an active area of research (see [1] for a recent survey of progress in this area). Most noticeably, the secrecy-capacity of the Gaussian wiretap channel (WTC) was characterized in [2]–[4] for the multiple-input single-output scenario, and in [5]–[9] for the MIMO case. The Gaussian MIMO WTC with a common message was studied in [10]. In [11], the secrecy-capacity region for the setting with a degraded message set and an external eavesdropper (from which all messages are concealed) was derived. The MIMO Gaussian BC with confidential messages, where the private message to each user is kept secret from the opposite user, without and with a common message was solved in [12] and [13], respectively. As the capacity region of the MIMO Gaussian BC without secrecy requirements was derived in [14] when no common message is present, and in [15] with a common message, this work settles the two remaining scenarios concerning secrecy. More specifically, focusing on the two-user MIMO Gaussian BC with or without a common message and where both, either or neither of the private messages are secret, we derive the secrecy-capacity regions of the only two instances that remained unsolved until now. A pointer to each past result and the contribution of this work are found in Table I.

Up until the more recent work of Geng and Nair [15], all the aforementioned results established the optimality of Gaussian inputs based on channel enhancement arguments, originally used in [14] to characterise the private message capacity region of the MIMO Gaussian BC (without secrecy constraints). In a nutshell, the idea of [14] was to approximate the actual BC using enhanced BCs, for which the entropy power inequality applies and is invoked to establish the optimality of Gaussian inputs (similarly to the proof for the scalar case by Bergmans [16]). Continuity arguments are then used to characterize the capacity region of the MIMO Gaussian BC of interest.

Adopting the approach of [15], in this work we take a more direct approach and prove the optimality of Gaussian inputs via factorization of upper concave envelopes (UCEs). We start by characterizing the secrecy-capacity region under an input covariance constraint for the setting with private and confidential messages only (i.e., when no common message is present). The derivation first describes the boundary points of a certain outer bound on the secrecy-capacity region as an UCE of a function of the input distribution. Having that, we show that if this UCE satisfies a specific factorization property, then it is maximized by a Gaussian input distribution. Then, using an

TABLE I: MIMO Gaussian BCs with/without a common message and where none/some/all of the private messages are secret - Summary of Results

$M_0$	$M_1$	$M_2$	Solution
—	Private	Private	Weingarten-Steinberg-Shamai 2006 [14]
Public	Private	Private	Geng-Nair 2014 [15]
Public	Secret from User 2	—	Ly-Liu-Liang 2010 [10]
—	Secret from User 2	Secret from User 1	Liu-Liu-Poor-Shamai 2010 [12]
Public	Secret from User 2	Secret from User 1	Ekrem-Ulukus 2012 [13]
—	Secret from User 2	Private	This work
Public	Secret from User 2	Private	This work

adaptation of dirty-paper coding (DPC) [17], we establish the equivalence of the outer bound to a particular inner bound, thus characterizing the secrecy-capacity region. Interestingly, optimality is achieved by using DPC to cancel out the signal of the non-confidential message  $M_2$  at Receiver 1 only. The other variant, i.e., DPC against the signal of the confidential message  $M_1$ , turns out to be unnecessary. This is in contrast to the case without secrecy requirements [15], where both variants of DPC are necessary to exhaust the entire region.

We then focus on the MIMO Gaussian BC with common, private and confidential messages (Fig. 1) and derive our main result by characterizing its secrecy-capacity region. Although this is a generalization of the problem without a common message, the secrecy-capacity of the latter setting is solved first. In doing so, we use the result without a common message to show that Gaussian inputs are optimal for a certain portion of the region with a common message. The rest of the region is characterized by extending the tools from [15] and introducing the notion of a double-nested UCE. Gaussian inputs once again are shown to exhaust the entire region. Finally, we visualize our results by a numerical example. Since the obtained regions are described as non-convex matrix optimization problems, we convert them into a computationally efficient form by relying on matrix decomposition properties from [18].

This paper is organized as follows. Section II gives definitions and describes the MIMO Gaussian BC with common, private and confidential messages. In Section III we state our main results: the secrecy-capacity regions of the considered BC without and with a common message. A numerical example that visualizes the obtained regions is also given in Section III. The mathematical background for proving our secrecy-capacity results is the focus of Section IV, where UCEs are studied. The proof of the results from Section III and the proofs of the properties from Section IV are given in Sections V and VI, respectively. Finally, Section VII summarizes the main achievements and insights of this paper.

## II. NOTATIONS AND PRELIMINARIES

We use the following notations. As customary  $\mathbb{N}$  is the set of natural numbers (which does not include 0), while  $\mathbb{R}$  are the reals. We further define  $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$ . Given two real numbers  $a, b$ , we denote by  $[a:b]$  the set of integers  $\{n \in \mathbb{N} | [a] \leq n \leq [b]\}$ . Calligraphic letters denote sets, e.g.,  $\mathcal{X}$ , while  $|\mathcal{X}|$  stands for the cardinality

of  $\mathcal{X}$ .  $\mathcal{X}^n$  denotes the  $n$ -fold Cartesian product of  $\mathcal{X}$ . An element of  $\mathcal{X}^n$  is denoted by  $x^n = (x_1, x_2, \dots, x_n)$ ; whenever the dimension  $n$  is clear from the context, vectors (or sequences) are denoted by boldface letters, e.g.,  $\mathbf{x}$ . The transpose of  $\mathbf{x}$  is denoted by  $\mathbf{x}^\top$ . Random variables are denoted by uppercase letters, e.g.,  $X$ , with similar conventions for random vectors. All the random variables considered in this work are real valued.

Probability distribution functions (PDFs) are denoted by the capital letters  $P$  or  $Q$ , with a subscript that identifies the random variable and its possible conditioning. For example, for two jointly continuous random vectors  $\mathbf{X}$  and  $\mathbf{Y}$ , let  $P_{\mathbf{X}}$ ,  $P_{\mathbf{X},\mathbf{Y}}$  and  $P_{\mathbf{X}|\mathbf{Y}}$  denote, respectively, the PDF of  $\mathbf{X}$ , the joint PDF of  $(\mathbf{X}, \mathbf{Y})$  and the conditional PDF of  $\mathbf{X}$  given  $\mathbf{Y}$ . Expressions such as  $P_{X,Y} = P_X P_{Y|X}$  are to be understood as  $P_{X,Y}(x, y) = P_X(x) P_{Y|X}(y|x)$ , for all  $(x, y) \in \mathcal{X} \times \mathcal{Y}$ . Accordingly, when three random variables  $X$ ,  $Y$  and  $Z$  satisfy  $P_{X|Y,Z} = P_{X|Y}$ , they form a Markov chain, which we denote by  $X - Y - Z$ . The subscripts of a PDF are omitted if its arguments are lowercase versions of the corresponding random variables. The expectation of a random variable  $X$  is denoted by  $\mathbb{E}X$ . When a random variable  $X$  is normally distributed we write  $X \sim \mathcal{N}(\mu, \sigma^2)$ , where  $\mu = \mathbb{E}X$  is the expectation of  $X$  and  $\sigma^2 = \text{var}(X)$  is its variance. Similarly, a vector Gaussian distribution of dimension  $n$  is defined by the expectation  $\boldsymbol{\mu} \in \mathbb{R}^n$  and the covariance matrix  $\mathbf{K} = \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top]$ . In general, non-italic capital letters, e.g.,  $\mathbf{A}$ , denote matrices. We use  $\mathbf{A} \succeq 0$  to indicate that a matrix  $\mathbf{A}$  is positive semidefinite, while  $\mathbf{A} \preceq \mathbf{B}$  denotes “less than or equal to” in the positive semidefinite ordering, i.e.,  $\mathbf{B} - \mathbf{A} \succeq 0$ .

**Definition 1 (Upper Concave Envelope)** *Let  $f : \mathcal{D} \rightarrow \mathbb{R}$  be a function defined on a convex set  $\mathcal{D}$ . The UCE  $\mathfrak{C}f : \mathcal{D} \rightarrow \mathbb{R}$  of  $f$  is the pointwise smallest concave function such that  $(\mathfrak{C}f)(x) \geq f(x)$ ,  $\forall x \in \mathcal{D}$ .*

Another representation of the UCE  $\mathfrak{C}f$  of  $f$  relies on the supporting hyperplanes of  $f$ . Namely, for any  $x \in \mathcal{D}$ , we have  $(\mathfrak{C}f)(x) = \sup_{V: \mathbb{E}[V]=x} \mathbb{E}[f(V)]$ .

#### A. Problem Definition

The outputs of a MIMO Gaussian BC at the  $i$ -th channel use are:

$$\mathbf{Y}_j(i) = \mathbf{G}_j \mathbf{X}(i) + \mathbf{Z}_j(i), \quad j = 1, 2, \quad i \in [1 : n], \quad (1)$$

where  $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{t \times t}$  are channel gain matrices (assumed to be known to all parties),  $\{\mathbf{Z}_j(i)\}_{i \in [1:n]}$ , for  $j = 1, 2$ , is an independent and identically distributed (i.i.d.) sequence of Gaussian random vectors taking values in  $\mathbb{R}^{t \times 1}$ . For each  $j = 1, 2$  and  $i \in [1 : n]$ , the elements of  $\mathbf{Z}_j(i) = [Z_{j,1}(i) \ Z_{j,2}(i) \ \dots \ Z_{j,t}(i)]^\top$  are also i.i.d. Gaussian random variables, whose expected values and variance are specified by the parameters of the normal distribution of  $\mathbf{Z}_j(i)$ . The input sequence  $\{\mathbf{X}(i)\}_{i \in [1:n]}$  is subject to the covariance constraint

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} [\mathbf{X}(i) \mathbf{X}^\top(i)] \preceq \mathbf{K}, \quad (2)$$

where  $\mathbf{K} \succeq 0$ .

**Remark 1** *We make the following assumptions on the channel gain matrices and the noise covariances:*

- 1)  $G_1$  and  $G_2$  are invertible. The analysis in this work relies on showing that certain inner and outer bounds on the secrecy-capacity region coincide. These bounds are characterized in terms of mutual informations between the channel input (or some auxiliary random variables) and the channel outputs. The mutual information terms, and hence the inner and outer bounds, are continuous functions of the channel gain matrices. Since the set of invertible matrices is a dense open set in the set of all  $t \times t$  matrices, by continuity of the bounds, the inner and outer bounds coincide for all channel gain matrices.
- 2) For each  $j = 1, 2$ , the Gaussian noise vectors,  $\{\mathbf{Z}_j(i)\}_{i \in [1:n]}$ , are i.i.d. according to  $\mathcal{N}(0, \mathbf{I})$ , where  $\mathbf{I}$  is the  $t \times t$ -identity matrix. This assumption is without loss of generality due to the following reasons: First, the mean of the Gaussian noise does not affect the capacity region. Second, when the covariance matrix is invertible, the noises can be whitened by multiplying (1) by another invertible matrix. On the other hand, if the covariance matrix is non-invertible, the communication scenario degenerates. This is since a suitable linear transformation converts the Gaussian channel to be noiseless, thus having an infinite capacity.

We study the scenario of a MIMO Gaussian BC with common, private and confidential messages (Fig. 1). The sender communicates three messages  $(M_0, M_1, M_2)$  over the MIMO Gaussian BC from (1).  $M_0$  is a common message that is intended to both users, while  $M_j$ , for  $j = 1, 2$ , is delivered to user  $j$  only. The receivers are to recover their intended messages with arbitrarily high probability. Moreover,  $M_1$  is a confidential message that is to be kept secret from User 2, which is formally described by the weak-secrecy requirement

$$\frac{1}{n} I(M_1; \mathbf{Y}_2^n) \xrightarrow{n \rightarrow \infty} 0, \quad (3)$$

where  $n$  is the number of channel uses. In (3),  $\mathbf{Y}_2^n \triangleq [\mathbf{Y}_2(1) \ \mathbf{Y}_2(2) \ \dots \ \mathbf{Y}_2(n)]^\top$ , where for each  $i \in [1 : n]$ ,  $\mathbf{Y}_2(i)$  is the output random vector (taking values in  $\mathbb{R}^t$ ) observed by User 2 at the  $i$ -th channel instance.

For any covariance constraint  $\mathbf{K} \succeq 0$ , the secrecy-capacity region  $\mathcal{C}_K$  is the closure of all achievable rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ , where achievability is defined in a standard manner (see, e.g., [19]).

### III. SECRECY-CAPACITY RESULTS

#### A. MIMO Gaussian BCs with Private and Confidential Messages

The MIMO Gaussian BC with private and confidential messages but without a common message is defined as in Section II-A, while setting  $R_0 = 0$ . For any covariance constraint  $\mathbf{K} \succeq 0$ , let  $\hat{\mathcal{C}}_K$  be the corresponding secrecy-capacity region, and for any  $0 \preceq \mathbf{K}^* \preceq \mathbf{K}$  set the following shorthand notations:

$$\hat{r}_1(\mathbf{K}^*) \triangleq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1 \mathbf{K}^* \mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_2 \mathbf{K}^* \mathbf{G}_2^\top} \right| \quad (4a)$$

$$\hat{r}_2(\mathbf{K}^*) \triangleq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2 \mathbf{K} \mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2 \mathbf{K}^* \mathbf{G}_2^\top} \right|. \quad (4b)$$

Define also

$$\hat{\mathcal{C}}_{\mathbf{K}}(\mathbf{K}^*) \triangleq \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \left| \begin{array}{l} R_1 \leq \hat{r}_1(\mathbf{K}^*) \\ R_2 \leq \hat{r}_2(\mathbf{K}^*) \end{array} \right. \right\}. \quad (5)$$

With that, the following theorem characterizes the secrecy-capacity region  $\hat{\mathcal{C}}_{\mathbf{K}}$ .

**Theorem 1 (Secrecy-Capacity without Common Message)** *The secrecy-capacity region  $\hat{\mathcal{C}}_{\mathbf{K}}$  of the MIMO Gaussian BC with private and confidential messages under the covariance constraint (2) is*

$$\hat{\mathcal{C}}_{\mathbf{K}} = \bigcup_{0 \preceq \mathbf{K}^* \preceq \mathbf{K}} \hat{\mathcal{C}}_{\mathbf{K}}(\mathbf{K}^*). \quad (6)$$

The proof of Theorem 1 (given in Section V-A) shows that certain inner and outer bounds of the secrecy-capacity region coincide, and that Gaussian inputs are optimal. First, we show that the boundary points of the outer bound are an UCE of a function of the input distribution. Based on some properties of UCEs (see Sections IV-A and IV-B) we deduce that a Gaussian input distribution maximizes the considered UCE. The secrecy-capacity region is then characterized by evaluating the boundary points of the inner bound under a Gaussian input vector and showing that they coincide with those of the outer bound.

**Remark 2 (Interpretation of Optimal Secrecy Rates)** *The right-hand side (RHS) of (4a) is the secrecy-capacity of the Gaussian MIMO WTC with input covariance  $\mathbf{K}^*$ , where User 1 serves as the legitimate party and User 2 as the eavesdropper. The RHS of (4b) is the capacity of the MIMO Gaussian point-to-point channel with input covariance  $\mathbf{K} - \mathbf{K}^*$  and noise covariance  $\mathbf{I} + \mathbf{K}^*$ . Thus,  $M_1$  being confidential forces User 1 to treat the second user as an eavesdropper. Then, the transmission rate of  $M_2$  (to User 2) is maximized by consuming the remaining power, while treating the signal of the first user as noise. The optimization over  $\mathbf{K}^*$  corresponds to different choices of user prioritization.*

**Remark 3 (Relation to Dirty-Paper Coding)** *As evident from the proof of Theorem 1 (see Proposition 8 in Section V-A), the entire secrecy-capacity region  $\hat{\mathcal{C}}_{\mathbf{K}}$  is achieved by using DPC to cancel out the signal of the non-confidential message  $M_2$  at Receiver 1 only. The other variant, i.e., DPC against the signal of the confidential message  $M_1$  at Receiver 2, is unnecessary. This is in contrast to when there is no secrecy requirement on  $M_1$  (namely, the private message BC), where the capacity region is exhausted by taking the convex hull of both variants (DPC against  $M_1$  and DPC against  $M_2$ ).*

**Remark 4 (Relation Common Message Case)** *Although being a special case of the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages  $\mathcal{C}_{\mathbf{K}}$ , the result of Theorem 1 is derived separately. This is since we use the fact that the boundary points of  $\hat{\mathcal{C}}_{\mathbf{K}}$  are achieved by Gaussian inputs to prove their optimality for a certain part  $\mathcal{C}_{\mathbf{K}}$ . More specifically, when the rate of the private message  $M_2$  is larger than this of the common message  $M_0$ , we show that characterizing  $\mathcal{C}_{\mathbf{K}}$  reduces to the characterization of  $\hat{\mathcal{C}}_{\mathbf{K}}$  and use Theorem 1.*

As a corollary of Theorem 1, we characterize the secrecy-capacity region under the average total power constraint. This is a simple consequence of [14, Lemma 1]; further details are omitted.

**Corollary 1 (Average Total Power Constraint)** *The secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages under the average total power constraint*

$$\frac{1}{n} \sum_{i=1}^n \|\mathbf{X}(i)\|^2 \leq P, \quad (7a)$$

is given by

$$\hat{\mathcal{C}}_P = \bigcup_{0 \preceq \mathbf{K}: \text{tr}(\mathbf{K}) \leq P} \hat{\mathcal{C}}_{\mathbf{K}}. \quad (7b)$$

**Remark 5 (Computing Secrecy-Capacity Region)** *In general, it is hard to compute (6) or (7b) as they involve non-convex matrix optimization problems. Nonetheless, in Section III-C we show how to convert (6) or (7a) into a computationally efficient form based on matrix decomposition properties from [18]. The simplified optimization problem is then used to illustrate the secrecy-capacity region under an average total power constraint  $\hat{\mathcal{C}}_P$  based on a numerical example.*

#### B. MIMO Gaussian BCs with Common, Private and Confidential Messages

To state the secrecy-capacity region  $\mathcal{C}_{\mathbf{K}}$  of the MIMO Gaussian BC with common, private and confidential messages as defined in Section II-A, we define

$$r_0^{(j)}(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_j \mathbf{K} \mathbf{G}_j^\top}{\mathbf{I} + \mathbf{G}_j (\mathbf{K}_1 + \mathbf{K}_2) \mathbf{G}_j^\top} \right|, \quad j = 1, 2 \quad (8a)$$

$$r_1(\mathbf{K}_2) = \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1 \mathbf{K}_2 \mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_2 \mathbf{K}_2 \mathbf{G}_2^\top} \right| \quad (8b)$$

$$r_2(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2 (\mathbf{K}_1 + \mathbf{K}_2) \mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2 \mathbf{K}_2 \mathbf{G}_2^\top} \right|, \quad (8c)$$

and set

$$\mathcal{C}_{\mathbf{K}}(\mathbf{K}_1, \mathbf{K}_2) \triangleq \left\{ (R_0, R_1, R_2) \in \mathbb{R}_+^3 \left| \begin{array}{l} R_0 \leq \min \{r_0^{(1)}(\mathbf{K}_1, \mathbf{K}_2), r_0^{(2)}(\mathbf{K}_1, \mathbf{K}_2)\} \\ R_1 \leq r_1(\mathbf{K}_2) \\ R_2 \leq r_2(\mathbf{K}_1, \mathbf{K}_2) \end{array} \right. \right\}. \quad (9)$$

**Theorem 2 (Secrecy-Capacity with Common Message)** *The secrecy-capacity region  $\mathcal{C}_{\mathbf{K}}$  of the MIMO Gaussian BC with common, private and confidential messages under the covariance constraint (2) is*

$$\mathcal{C}_{\mathbf{K}} = \bigcup_{\substack{0 \preceq \mathbf{K}_1, \mathbf{K}_2: \\ \mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{K}}} \mathcal{C}_{\mathbf{K}}(\mathbf{K}_1, \mathbf{K}_2). \quad (10)$$

Theorem 2 is proven in Section V-B.

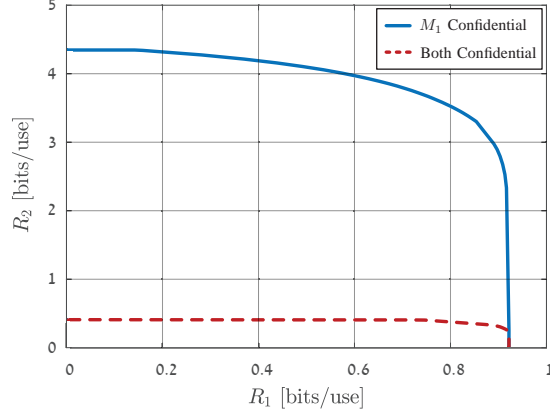


Fig. 2: Secrecy-capacity region under an average total power constraint of the MIMO Gaussian BC without a common message, where:  $M_1$  is confidential and  $M_2$  is private (solid blue) vs.  $M_1$  and  $M_2$  are both confidential (dashed red).

**Remark 6 (Interpretation of Optimal Secrecy Rates)** *Our interpretation of the structure of  $\mathcal{C}_K$  is reminiscent of Remark 2. First, (8b) indicates that User 1 achieves rates up to the secrecy-capacity of the MIMO Gaussian WTC with input covariance  $K_2$ . The 2nd user treat this signal as an additive Gaussian noise when decoding its private message  $M_2$ , which is transmitted using another (independent) Gaussian signal with covariance  $K_1$  (see (8c)). According to (8a), the remaining portion of the total covariance matrix, that is,  $K - (K_1 + K_2)$ , is employed to encode the common message  $M_0$ , which is decoded by each receiver while treating all other signals as noise. As in the case without a common message, a layered coding scheme, when optimized over the choices of  $K_1$  and  $K_2$ , exhausts the entire secrecy-capacity region.*

As before, Theorem 2 gives rise to a characterization of the secrecy-capacity region under the average total power constraint.

**Corollary 2 (Average Total Power Constraint)** *The secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages under the average total power constraint (7a) is given by*

$$\mathcal{C}_P = \bigcup_{0 \preceq K: \text{tr}(K) \leq P} \mathcal{C}_K. \quad (11)$$

### C. Numerical Example

We illustrate the secrecy-capacity region  $\hat{\mathcal{C}}_P$  of the MIMO Gaussian BC with private and confidential messages (without a common message) under an average total power constraint  $P$  (Corollary 1). The region is described in (7b) as the union of all secrecy-capacity regions  $\hat{\mathcal{C}}_K$ , each under a covariance constraint  $K$  with  $\text{tr}(K) \leq P$ . However,  $\hat{\mathcal{C}}_K$  itself is described as matrix optimization problems that is not convex in general, and is therefore, hard to compute.



We overcome the computational inefficiency of  $\hat{\mathcal{C}}_K$  by leveraging the decomposition proposed in [18, Equation (10)]: Every positive semidefinite matrix  $K^* \in \mathbb{R}^{t \times t}$  with  $K^* \preceq K$  can be expressed as

$$K^* = K^{\frac{1}{2}} V D V^\top K^{\frac{1}{2}}, \quad (12)$$

where  $V \in \mathbb{R}^{t \times t}$  is a unitary matrix and  $D \in \mathbb{R}^{t \times t}$  is a diagonal matrix whose diagonal values are between 0 and 1. Since in the subsequent example the dimension is  $t = 2$ , a general unitary matrix  $V$  is just a rotation matrix, i.e., we set

$$V = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \theta \in [0, 2\pi]. \quad (13)$$

Running over all possible diagonal matrices  $D$  involves only two parameters, viz. the diagonal entries of  $D$ . Finally, note that  $K^{\frac{1}{2}}$  is any matrix  $B$  satisfying  $BB^\top = K$ . Obviously, there are many such matrices (in fact if  $B$  satisfies  $BB^\top = K$ , then so does  $BU$ , for any unitary  $U$ ). However, since the numerical calculation runs over all matrices  $V$  from (13) anyway, any choice of  $B$  would do. Our simulation uses the Cholesky decomposition of  $K$  to calculate  $B$ .

The region  $\hat{\mathcal{C}}_P$  is computed according to (7b), while noting that one may restrict the optimization domain to positive semidefinite matrices  $K$  with  $\text{tr}(K) = P$ . This observation follows because for every  $K'$  with  $\text{tr}(K') = \pi < P$ , there is a  $K$  with  $\text{tr}(K) = P$ , such that

$$\hat{\mathcal{C}}_{K'} \subseteq \hat{\mathcal{C}}_K. \quad (14)$$

The matrix  $K$  is constructed by increasing the  $(1, 1)$ -th entry of  $K'$  by  $P - \pi$ , while all other entries of  $K'$  remain unchanged. The construction satisfied  $K' \preceq K$  and the inclusion in (14) follows because fixing  $K^* \preceq K' \preceq K$  and replacing  $K'$  with  $K$  in (4) does not alter (4a) and strictly increases (4b).

In the numerical example we set

$$G_1 = \begin{bmatrix} 0.3 & 2.5 \\ 2.2 & 1.8 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 1.3 & 1.2 \\ 1.5 & 3.9 \end{bmatrix} \quad (15)$$

and  $P = 12$ . The secrecy-capacity region  $\hat{\mathcal{C}}_P$  is given by the solid blue curve in Fig. 2. For comparison, the secrecy-capacity region of the MIMO Gaussian BC with confidential messages [12] (i.e., when each user serves as the eavesdropper of the message to the other user) is depicted by the dashed red curve. Fig. 2 shows that imposing a secrecy constraint on  $M_2$  at the 1st receiver strictly shrinks the secrecy-capacity region. Although in *both* regions the maximal value of  $R_1$  is the secrecy-capacity of the corresponding MIMO Gaussian WTC (see (4a) and [12, Equation (4)]), the achievable values of  $R_2$  drop if  $M_2$  is also confidential.

#### IV. OPTIMALITY OF GAUSSIAN INPUTS VIA FACTORIZATION OF CONCAVE ENVELOPES

This section provides the mathematical background for characterizing the secrecy-capacity regions of the considered MIMO Gaussian BC without and with a common message (Theorems 1 and 2). More specifically, we define some generic functions and show that they are maximized by Gaussian distributions. These functions are subsequently used to describe the boundary points of certain outer bounds on the secrecy-capacity regions of interest. The properties established in this section are then used to show that optimality is achieved by Gaussian inputs, and that the resulting expressions can be achieved by a corresponding inner bound.

Sections IV-A and IV-B focus on functions that are reminiscent of those studied in [15, Sections II-B and II-C]. Therefore, to avoid verbatim repetition of arguments from [15], we state some of the properties in Sections IV-A and IV-B without proofs. The focus of Section IV-C is on a new function that was not considered in [15], the properties of which we prove in full detail. All the proofs for this section are relegated to Section VI.

Establishing Gaussian inputs as maximizers relies on the notion two-letter BCs [15, Section I-A], which is a special case of a product BC (PBC)<sup>1</sup>.

**Definition 2 (Product BC)** *A PBC consists of a sender  $(\mathbf{X}_1, \mathbf{X}_2)$  and two receivers  $(\mathbf{Y}_{11}, \mathbf{Y}_{12})$  and  $(\mathbf{Y}_{21}, \mathbf{Y}_{22})$ , and is described by a transition probability of the form  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21} | \mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22} | \mathbf{X}_2}^{(2)}$ .*

A MIMO Gaussian PBC can be represented as

$$\begin{bmatrix} \mathbf{Y}_{11} \\ \mathbf{Y}_{12} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{11} & 0 \\ 0 & \mathbf{G}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{11} \\ \mathbf{Z}_{12} \end{bmatrix} \quad (16a)$$

$$\begin{bmatrix} \mathbf{Y}_{21} \\ \mathbf{Y}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{21} & 0 \\ 0 & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{21} \\ \mathbf{Z}_{22} \end{bmatrix}, \quad (16b)$$

where  $\mathbf{Z}_{11}, \mathbf{Z}_{12}, \mathbf{Z}_{21}, \mathbf{Z}_{22} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  are i.i.d. and independent of  $(\mathbf{X}_1, \mathbf{X}_2)$ . A *two-letter* version of a BC is a PBC where the components are identical, i.e.,  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21} | \mathbf{X}_1}^{(1)} = Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22} | \mathbf{X}_2}^{(2)}$ . In all subsequent definitions and results, the input covariance constraining matrix  $\mathbf{K} \succeq 0$  (see (2)) stays fixed.

##### A. Difference of Mutual Information Terms

Consider a BC  $Q_{\mathbf{Y}_1, \mathbf{Y}_2 | \mathbf{X}}$ . For any  $\eta > 1$ , let  $s_\eta^Q$  be a function of  $\mathbf{X} \sim P_{\mathbf{X}}$  defined by

$$s_\eta^Q(\mathbf{X}) \triangleq I(\mathbf{X}; \mathbf{Y}_2) - \eta I(\mathbf{X}; \mathbf{Y}_1). \quad (17)$$

**Remark 7** *The definition of  $s_\eta^Q(\mathbf{X})$  in (17) differs from the function  $s_\lambda^q(\mathbf{X}) \triangleq I(\mathbf{X}; \mathbf{Y}_1) - \lambda I(\mathbf{X}; \mathbf{Y}_2)$  defined in [15, Section II-B] only in the ordering of the mutual information terms and the labeling of the parameter. As none of these differences plays a role in deriving the properties of  $s_\lambda^q(\mathbf{X})$  given in [15], we restate and use some of these*

<sup>1</sup>Henceforth, we omit the time index  $i$ .

properties with respect to  $s_\eta^Q(\mathbf{X})$  without providing proofs. Additional attributes of  $s_\eta^Q(\mathbf{X})$  that were not established in [15] are rigorously derived.

For a pair of random variables  $(V, \mathbf{X})$  such that  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, set

$$s_\eta^Q(\mathbf{X}|V) \triangleq I(\mathbf{X}; \mathbf{Y}_2|V) - \eta I(\mathbf{X}; \mathbf{Y}_1|V), \quad (18)$$

and define the UCE of  $s_\eta^Q(\mathbf{X})$  as

$$S_\eta^Q(\mathbf{X}) \triangleq \left( \mathfrak{E} s_\eta^Q \right)(\mathbf{X}) = \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} s_\eta^Q(\mathbf{X}|V). \quad (19)$$

The second equality in (19) follows directly from Definition 1. For any discrete random variable  $V$  we also set  $S_\eta^Q(\mathbf{X}|V) \triangleq \sum_v P(v) S_\eta^Q(\mathbf{X}|V = v)$ , and naturally extend this definition for an arbitrary  $V$ .

**Proposition 1 (Concave Envelopes Properties)** *The UCE  $S_\eta^Q$  satisfies:*

- 1) *If  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, then  $S_\eta^Q(\mathbf{X}|V) \leq S_\eta^Q(\mathbf{X})$ .*
- 2) *If  $W - V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, then  $S_\eta^Q(\mathbf{X}|V, W) = S_\eta^Q(\mathbf{X}|V)$ .*
- 3)  *$S_\eta^Q(\mathbf{X})$  is convex in  $\eta$  inside  $(0, 2)$ , for a fixed  $P_{\mathbf{X}}$ , and therefore it is continuous in  $\eta$  at  $\eta = 1$  <sup>2</sup>.*

The proof of Proposition 1 is relegated to Section VI-A.

**Definition 3 (Maximized Concave Envelope)** *For any MIMO Gaussian BC  $Q_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}}$ , define*

$$V_\eta^Q(K) \triangleq \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} S_\eta^Q(\mathbf{X}) = \sup_{\substack{(V, \mathbf{X}): \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} s_\eta^Q(\mathbf{X}|V). \quad (20)$$

We subsequently show that (20) is achieved by a Gaussian input distribution. At the heart of the proof is a specific factorization property of the UCE  $S_\eta^Q$ . To formulate this property, we first extend  $S_\eta^Q$  to PBCs. For a PBC  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$  we set,

$$s_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) \triangleq I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) - \eta I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}), \quad (21)$$

and define the quantities  $s_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ ,  $S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2)$  and  $S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$  analogously to the definitions of  $s_\eta^Q(\mathbf{X}|V)$ ,  $S_\eta^Q(\mathbf{X})$  and  $S_\eta^Q(\mathbf{X}|V)$  from the above, respectively.

**Proposition 2 (Factorization Property)** *For any PBC  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$ , the following chain of inequalities holds*

$$S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) \leq S_\eta^{Q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + S_\eta^{Q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) \leq S_\eta^{Q_1}(\mathbf{X}_1) + S_\eta^{Q_2}(\mathbf{X}_2). \quad (22)$$

<sup>2</sup>The crux of the 3rd property is the continuity of the UCE in  $\eta$  at  $\eta = 1$ , which can be established by considering any bounded, open interval containing 1, and not necessarily  $(0, 2)$ .

The proof of Proposition 2 follows by repeating the steps in the proof of [15, Proposition 6], while switching the roles of  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$ . We omit the proof due the similarity.

**Theorem 3 (Existence and Uniqueness of Gaussian Maximizer)** *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$ . There exists a decomposition  $\mathbf{X} = \mathbf{X}^* + \mathbf{X}'$ , such that  $\mathbf{X}^*$  and  $\mathbf{X}'$  are independent,  $\mathbf{X}^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^*)$ ,  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathbf{K} - \mathbf{K}^*)$ , where  $\mathbf{K}^* \preceq \mathbf{K}$ , and  $S_\eta^Q(\mathbf{X}) = s_\eta^Q(\mathbf{X}^*) = V_\eta^Q(\mathbf{K})$ . Furthermore, this decomposition (i.e., the covariance matrix  $\mathbf{K}^*$ ) is unique.*

The proof of Theorem 3 is also omitted as it mimics the proofs of Theorem 1 and Corollary 1 in [15].

### B. Nested Upper Concave Envelopes

The function considered in this subsection is used to derive the secrecy-capacity region of the considered MIMO Gaussian BC without a common message (see Section V-A). The result for the case with a common message, which is the main focus of this work, relies on this derivation by establishing an equivalence between a certain portion of secrecy-capacity region with a common message and the region when no common message is present.

For a BC  $Q_{\mathbf{Y}_1, \mathbf{Y}_2 | \mathbf{X}}$ ,  $\eta > 1$ ,  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ , where  $\lambda_j > 0$ ,  $j = 1, 2$ , and any  $\mathbf{X} \sim P_{\mathbf{X}}$  define

$$t_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}) \triangleq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2) + \lambda_1 S_\eta^Q(\mathbf{X}), \quad (23)$$

where  $S_\eta^Q(\mathbf{X})$  is given by (19). As before, for a pair of random variables  $(V, \mathbf{X})$  for which  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, let

$$t_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V) \triangleq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2|V) + \lambda_1 S_\eta^Q(\mathbf{X}|V), \quad (24)$$

and set

$$T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}) \triangleq \mathfrak{C}(t_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X})) = \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} t_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V). \quad (25)$$

Define  $T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V) \triangleq \sum_v P(v) T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V = v)$ , for a  $V$  with a countable alphabet and consider its natural extension when  $V$  is arbitrary.

**Remark 8 (Nested Concave Envelopes Properties)** *Similarly to the properties of  $S_\eta^Q$  stated in Proposition 1, since  $T_{\boldsymbol{\lambda}, \eta}^Q$  is concave in  $P_{\mathbf{X}}$ , Jensen's inequality implies that  $T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V) \leq T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X})$ , for any  $(V, \mathbf{X})$  satisfying  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ . Moreover, if  $W - V - \mathbf{X}$  forms a Markov chain, then  $T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|W, V) = T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V)$ , because  $P_{\mathbf{X}|W, V} = P_{\mathbf{X}|V}$ . Finally,  $T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X})$  is convex in  $\eta$  inside  $(0, 2)$ , for a fixed  $P_{\mathbf{X}}$ , and therefore it is continuous in  $\eta$  at  $\eta = 1$ .*

**Definition 4 (Maximized Nested Concave Envelope)** *For any MIMO Gaussian BC  $Q_{\mathbf{Y}_1, \mathbf{Y}_2 | \mathbf{X}}$ , define*

$$\hat{V}_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{K}) \triangleq \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^T] \preceq \mathbf{K}} T_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}) = \sup_{\substack{(V, \mathbf{X}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] \preceq \mathbf{K}, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} t_{\boldsymbol{\lambda}, \eta}^Q(\mathbf{X}|V). \quad (26)$$

**Proposition 3 (Continuity of Maximal Value)** *For any  $\lambda$  as before,  $\hat{V}_{\lambda,\eta}(\mathbf{K})$  is continuous in  $\eta$  at  $\eta = 1$ .*

The proof of Proposition 3 follows by arguments similar to those in the proof of Property 3 of Proposition 1. Namely, the continuity of  $\hat{V}_{\lambda,\eta}(\mathbf{K})$  at  $\eta = 1$  is established by verifying that  $\hat{V}_{\lambda,\eta}(\mathbf{K})$  is convex in  $\eta$  inside  $(0, 2)$  and using Proposition 17 from [20, Chapter 5].

As before, to state the factorization property for nested UCEs, we extend some of the preceding definitions to PBCs. Let  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$  be a PBC and we set,

$$t_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) = \lambda_1 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) + \lambda_1 S_{\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2). \quad (27)$$

Furthermore, define  $t_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ ,  $T_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2)$  and  $T_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$  in a similar manner to  $t_{\lambda,\eta}^Q(\mathbf{X}|V)$ ,  $T_{\lambda,\eta}^Q(\mathbf{X})$  and  $T_{\lambda,\eta}^Q(\mathbf{X}|V)$ , respectively. The following proposition states the  $T_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$  factorization property on interest, which subsequently serves as the key ingredient in proving the existence of a Gaussian maximizer for  $\hat{V}_{\lambda,\eta}^Q(\mathbf{K})$  from (26).

**Proposition 4 (Factorization Property)** *For any PBC  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$ , the following chain of inequalities holds*

$$T_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) \leq T_{\lambda,\eta}^{Q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + T_{\lambda,\eta}^{Q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) \leq T_{\lambda,\eta}^{Q_1}(\mathbf{X}_1) + T_{\lambda,\eta}^{Q_2}(\mathbf{X}_2) \quad (28)$$

Furthermore, if the PBC is Gaussian and a triple  $(V^*, \mathbf{X}_1^*, \mathbf{X}_2^*)$  satisfies

$$t_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1^*, \mathbf{X}_2^*|V^*) = T_{\lambda,\eta}^{Q_1 \times Q_2}(\mathbf{X}_1^*, \mathbf{X}_2^*) = T_{\lambda,\eta}^{Q_1}(\mathbf{X}_1^*) + T_{\lambda,\eta}^{Q_2}(\mathbf{X}_2^*), \quad (29)$$

then  $\mathbf{X}_1^* - V^* - \mathbf{X}_2^*$  and  $t_{\lambda,\eta}^{Q_j}(\mathbf{X}_j^*|V^*) = T_{\lambda,\eta}^{Q_j}(\mathbf{X}_j^*)$ , for  $j = 1, 2$ .

See Section VI-B for the proof of Proposition 4. Having this, the existence of a Gaussian maximizer for  $\hat{V}_{\lambda,\eta}^Q(\mathbf{K})$  follows by repeating the proofs of Theorem 2 and Corollary 2 in [15] with respect to our definition of  $T_{\lambda,\eta}^Q$ . The existence is stated in the following Theorem, which we give without proof.

**Theorem 4 (Existence and Uniqueness of Gaussian Maximizer)** *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$ . There exists a unique decomposition  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}'$  into independent random variables  $(\mathbf{X}_1^*, \mathbf{X}_2^*, \mathbf{X}')$ , where  $\mathbf{X}_j^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_j)$ ,  $j = 1, 2$ , and  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathbf{K} - (\mathbf{K}_1 + \mathbf{K}_2))$ ,  $\mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{K}$ , such that*

$$T_{\lambda,\eta}^Q(\mathbf{X}) = t_{\lambda,\eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = \hat{V}_{\lambda,\eta}^Q(\mathbf{K}) \quad (30a)$$

$$S_{\eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = s_{\eta}^Q(\mathbf{X}_1^*) = V_{\eta}^Q(\mathbf{K}_1 + \mathbf{K}_2). \quad (30b)$$

### C. Double-Nested Upper Concave Envelopes

The definitions and properties in this section are used to derive the secrecy-capacity region of the cooperative BC with common, private and confidential messages (as defined in Section II-A). Let  $Q_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}}$ ,  $\eta > 1$  be a BC,

$\lambda_0 = (\lambda_0, \lambda_1, \lambda_2)$ , where  $\lambda_j > 0$  for  $j = 0, 1, 2$  and  $\lambda_0 > \lambda_2$ ,  $\alpha \in [0, 1]$  and  $\bar{\alpha} = 1 - \alpha$ . For any  $\mathbf{X} \sim P_{\mathbf{X}}$  define

$$f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) \triangleq (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + T_{\lambda, \eta}^Q(\mathbf{X}), \quad (31)$$

where  $T_{\lambda, \eta}^Q(\mathbf{X})$  is given by (25) and  $\lambda = (\lambda_1, \lambda_2)$ .

For  $(V, \mathbf{X})$  that satisfy the Markov chain  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ , we set  $f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V)$  in an analogous manner to (24), while  $F_{\lambda_0, \alpha, \eta}^Q \triangleq \mathfrak{C}f_{\lambda_0, \alpha, \eta}^Q$  denotes the UCE of  $f_{\lambda_0, \alpha, \eta}^Q$ . We also set  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V) \triangleq \sum_v P(v)F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V = v)$  for a discrete  $V$  and, as before, consider its natural extension in the case where  $V$  is arbitrary.

**Remark 9 (Double-Nested Concave Envelopes Properties)** *The concavity of  $F_{\lambda_0, \alpha, \eta}^Q$  in  $P_{\mathbf{X}}$  and Jensen's inequality imply that for any  $(V, \mathbf{X})$  with  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ , it holds that  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V) \leq F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X})$ . If the chain  $W - V - \mathbf{X}$  is Markov, then  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|W, V) = F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V)$ , because  $P_{\mathbf{X}|W, V} = P_{\mathbf{X}|V}$ . As a function of  $\eta$ ,  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X})$  is convex inside  $(0, 2)$ , for any fixed  $\mathbf{X}$ , and is therefore continuous at  $\eta = 1$ .*

**Definition 5 (Maximized Double-Nested Concave Envelope)** *For any MIMO Gaussian BC  $Q_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}}$ , define*

$$\tilde{V}_{\lambda_0, \alpha, \eta}^Q(K) \triangleq \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) = \sup_{\substack{(V, \mathbf{X}): \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V). \quad (32)$$

**Remark 10 (Continuity of Maximal Value)** *As before, one can readily verify that as a function of  $\eta$ ,  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(K)$  is convex inside  $(0, 2)$ , and deduce its continuity at  $\eta = 1$ .*

The above notions are once again extended to PBC. Namely, for any PBC  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$ , we set

$$f_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) = (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) - \alpha\lambda_0 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) + T_{\lambda, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2), \quad (33)$$

and define  $f_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ ,  $F_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2)$  and  $F_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$  as the natural extensions to the PBC scenario of  $f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V)$ ,  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X})$  and  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V)$  given above, respectively.

Moving forward, the factorization property of  $F_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}$  is stated in Proposition 5, while Proposition 6 established the existence of its maximizer.

**Proposition 5 (Factorization Property)** *For any PBC  $Q_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times Q_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$ , the following chain of inequalities holds*

$$F_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2) \leq F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) \leq F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2) \quad (34)$$

*Furthermore, if the PBC is Gaussian and a triple  $(V^*, \mathbf{X}_1^*, \mathbf{X}_2^*)$  satisfies*

$$f_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1^*, \mathbf{X}_2^*|V^*) = F_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1^*, \mathbf{X}_2^*) = F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1^*) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2^*), \quad (35)$$

*then  $\mathbf{X}_1^* - V^* - \mathbf{X}_2^*$  and  $f_{\lambda_0, \alpha, \eta}^{Q_j}(\mathbf{X}_j^*|V^*) = F_{\lambda_0, \alpha, \eta}^{Q_j}(\mathbf{X}_j^*)$ , for  $j = 1, 2$ .*

See Section VI-C for the proof of Proposition 5.

**Proposition 6 (Existence of a Maximizer)** *There exists a pair  $(V^*, \mathbf{X}^*)$  with  $|\mathcal{V}^*| \leq \frac{t(t+1)}{2} + 1$  and  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}$ , such that*

$$\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_2^* | V^*). \quad (36)$$

Furthermore, one may assume that  $\mathbb{E}[\mathbf{X}^* | V^* = v^*] = \mathbf{0}$ , for every  $v^* \in \mathcal{V}^*$ .

The existence of a maximizer and the cardinality bound on  $\mathcal{V}^*$  are proven in Appendix B. A zero conditional expectation can be assumed because centering conditioned on each  $V^* = v^*$  does not change the mutual information terms and hence  $f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_2^* | V^*)$  remains unchanged as well. Also, the centered versions of the input continues to satisfy the covariance constraint.

To show that the distribution that achieves  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$  is Gaussian, an invariance of  $f_{\lambda_0, \alpha, \eta}^Q$  with respect to the rotation operation is required. This invariance property is stated in the following Proposition and proven in Section VI-D.

**Proposition 7 (Invariance to Rotation)** *Let  $(V, \mathbf{X}) \sim P_{V, \mathbf{X}}^*$  attain  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ , with  $|\mathcal{V}| = m \leq \frac{t(t+1)}{2} + 1$ , and let  $\mathbf{X}_v$  be a centered random variable (zero mean) distributed according to the conditional PMF  $P_{\mathbf{X}|V=v}^*$ . Let  $(V_1, \mathbf{X}_1, V_2, \mathbf{X}_2) \sim P_{V, \mathbf{X}}^* \times P_{V, \mathbf{X}}^*$  be two i.i.d. copies of  $(V^*, \mathbf{X}^*)$ . Define*

$$\begin{aligned} \tilde{V} &= (V_1, V_2) \\ \mathbf{X}_{\theta_1} | \{\tilde{V} = (v_1, v_2)\} &\sim \frac{1}{\sqrt{2}}(\mathbf{X}_{v_1} + \mathbf{X}_{v_2}) \\ \mathbf{X}_{\theta_2} | \{\tilde{V} = (v_1, v_2)\} &\sim \frac{1}{\sqrt{2}}(\mathbf{X}_{v_1} - \mathbf{X}_{v_2}), \end{aligned}$$

where  $\mathbf{X}_{v_1}$  and  $\mathbf{X}_{v_2}$  are taken to be independent random variables, i.e.,  $(\mathbf{X}_{v_1}, \mathbf{X}_{v_2}) \sim P_{\mathbf{X}|V=v_1}^* \times P_{\mathbf{X}|V=v_2}^*$ . Then  $\mathbf{X}_{\theta_1} - \tilde{V} - \mathbf{X}_{\theta_2}$  and  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_j} | \tilde{V})$ , for  $j = 1, 2$ .

The existence of a Gaussian Maximizer for  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$  is stated next.

**Theorem 5 (Existence of Gaussian Maximizer)** *There exists an  $\mathbf{X}^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^*)$ , where  $\mathbf{K}^* \preceq \mathbf{K}$ , such that  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = \tilde{h}_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*)$ . Furthermore, the zero mean maximizer is unique.*

See Section VI-E for the proof.

**Corollary 3 (Gaussian Maximizer Properties)** *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$ . There is a unique decomposition  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^* + \mathbf{X}'$  into independent random variables  $(\mathbf{X}_1^*, \mathbf{X}_2^*, \mathbf{X}_3^*, \mathbf{X}')$ , where  $\mathbf{X}_j^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_j)$ ,  $j = 1, 2, 3$ , and  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathbf{K} - (\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3))$ ,  $\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3 \preceq \mathbf{K}$ , such that*

$$F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^*) = \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) \quad (37a)$$

$$T_{\lambda, \eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^*) = t_{\lambda, \eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = \hat{V}_{\lambda, \eta}^Q(\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3) \quad (37b)$$

$$S_{\eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = s_{\eta}^Q(\mathbf{X}_1^*) = V_{\eta}^Q(\mathbf{K}_1 + \mathbf{K}_2). \quad (37c)$$

Corollary 3, which is a consequence of Theorem 5, is our main tool in characterizing the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages. The proof of the Corollary is relegated to Section VI-F.

## V. PROOFS OF SECRECY-CAPACITY RESULTS

### A. Proof of Theorem 1

We establish the secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages by showing that certain outer bound and inner bounds match. In particular, we consider special cases of the inner and outer bounds from Theorems 1 and 2 of [21], respectively. To state the bounds, let  $\hat{\mathcal{C}}$  denote the secrecy-capacity region of the corresponding discrete-memoryless (DM) BC.

**Bound 1 (Outer Bound)** Let  $\hat{\mathcal{O}}$  be the closure of the union of rate pairs  $(R_1, R_2) \in \mathbb{R}_+^2$  satisfying:

$$R_1 \leq I(U; Y_1|V) - I(U; Y_2|V) \quad (38a)$$

$$R_2 \leq I(V; Y_2) \quad (38b)$$

over all  $(V, U) - X - (Y_1, Y_2)$ . Then  $\hat{\mathcal{C}} \subseteq \hat{\mathcal{O}}$ .

**Bound 2 (Inner Bound)** Let  $\hat{\mathcal{I}}$  be the closure of the union of rate pairs  $(R_1, R_2) \in \mathbb{R}_+^2$  satisfying:

$$R_1 \leq I(U; Y_1) - I(U; V) - I(U; Y_2|V) \quad (39a)$$

$$R_2 \leq I(V; Y_2) \quad (39b)$$

over all  $(V, U) - X - (Y_1, Y_2)$ . Then  $\hat{\mathcal{I}} \subseteq \hat{\mathcal{C}}$ .

The reader is referred to Appendix A for the proofs of Bounds 1 and 2. Let  $\hat{\mathcal{C}}_K$ ,  $\hat{\mathcal{O}}_K$  and  $\hat{\mathcal{I}}_K$  denote secrecy-capacity region, the outer bound and the inner bound for a MIMO Gaussian BC computed under a covariance input constraint  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}$ . Accordingly, we have  $\hat{\mathcal{I}}_K \subseteq \hat{\mathcal{C}}_K \subseteq \hat{\mathcal{O}}_K$ .

The opposite inclusion, i.e.,  $\hat{\mathcal{O}}_K \subseteq \hat{\mathcal{I}}_K$ , is shown next. The regions  $\hat{\mathcal{I}}_K$  and  $\hat{\mathcal{O}}_K$  are closed, convex and bounded subsets of the first quadrant, and therefore, are characterised by the intersection of their supporting hyperplanes.

**Lemma 1 (Supporting Hyperplanes)** The following are supporting hyperplanes of  $\hat{\mathcal{O}}_K$  and  $\hat{\mathcal{I}}_K$ :

$$R_1 \geq 0 \quad , \quad R_1 \leq \mathcal{H}_1^K \quad , \quad R_2 \geq 0 \quad , \quad R_2 \leq \mathcal{H}_2^K, \quad (40)$$

where

$$\mathcal{H}_1^K \triangleq \max_{(V,U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}} I(U; \mathbf{Y}_1|V) - I(U; \mathbf{Y}_2|V) \quad (41a)$$



$$\mathcal{H}_2^K \triangleq \max_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}} I(\mathbf{X}; \mathbf{Y}_2). \quad (41b)$$

Furthermore,  $(\mathcal{H}_1^K, 0)$  and  $(0, \mathcal{H}_2^K)$  are boundary points of  $\hat{\mathcal{O}}_K$  and  $\hat{\mathcal{I}}_K$ .

*Proof:* Clearly, (40) are supporting hyperplanes of  $\hat{\mathcal{O}}_K$  and the points  $(\mathcal{H}_1^K, 0)$  and  $(0, \mathcal{H}_2^K)$  are on its boundary. Furthermore,  $R_1 \geq 0$ ,  $R_2 \geq 0$  and  $R_2 \leq \mathcal{H}_2^K$  are also supporting hyperplanes of  $\hat{\mathcal{I}}_K$ , and since  $(0, \mathcal{H}_2^K) \in \hat{\mathcal{I}}_K$ , it is a boundary point of  $\hat{\mathcal{I}}_K$ . Note that  $\mathcal{H}_1^K$  describes the secrecy-capacity of the MIMO Gaussian WTC, where user 1 serves as the legitimate receiver and user 2 serves as the eavesdropper. Therefore (see [6]–[8]),

$$\mathcal{H}_1^K = \frac{1}{2} \max_{0 \preceq \mathbf{K}^* \preceq \mathbf{K}} \log |\mathbf{I} + \mathbf{G}_1 \mathbf{K}^* \mathbf{G}_1^\top| - \log |\mathbf{I} + \mathbf{G}_2 \mathbf{K}^* \mathbf{G}_2^\top|. \quad (42)$$

To see that  $(\mathcal{H}_1^K, 0)$  is also in  $\hat{\mathcal{I}}_K$  consider the following. For every  $0 \preceq \mathbf{K}^* \preceq \mathbf{K}$ , let  $\mathbf{X}_1$  and  $\mathbf{X}_2$  be independent Gaussian random vectors with covariances  $\mathbf{K}^*$  and  $\mathbf{K} - \mathbf{K}^*$ , respectively. Set

$$\mathbf{U} = \mathbf{X}_1 + \mathbf{A}\mathbf{X}_2$$

$$\mathbf{V} = \mathbf{X}_2$$

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2,$$

where  $\mathbf{A} = \mathbf{K}^* \mathbf{G}_1^\top [\mathbf{I} + \mathbf{G}_1 \mathbf{K}^* \mathbf{G}_1^\top]^{-1}$  is the precoding matrix for suppressing  $\mathbf{V}$  from  $\mathbf{Y}_1$  [17, Theorem 1]. Evaluating the mutual information terms on the RHS of (39a), we first have

$$I(\mathbf{U}; \mathbf{Y}_1) - I(\mathbf{U}; \mathbf{V}) = I(\mathbf{X}; \mathbf{Y}_1 | \mathbf{V}) = \frac{1}{2} \log |\mathbf{I} + \mathbf{G}_1 \mathbf{K}^* \mathbf{G}_1^\top|. \quad (43)$$

Moreover,

$$\begin{aligned} I(\mathbf{U}; \mathbf{Y}_2 | \mathbf{V}) &= I(\mathbf{X}_1 + \mathbf{A}\mathbf{X}_2; \mathbf{G}_2(\mathbf{X}_1 + \mathbf{X}_2) + \mathbf{Z}_2 | \mathbf{X}_2) \\ &= I(\mathbf{X}_1; \mathbf{G}_2 \mathbf{X}_1 + \mathbf{Z}_2 | \mathbf{X}_2) \\ &\stackrel{(a)}{=} I(\mathbf{X}_1; \mathbf{G}_2 \mathbf{X}_1 + \mathbf{Z}_2) \\ &= \frac{1}{2} \log |\mathbf{I} + \mathbf{G}_2 \mathbf{K}^* \mathbf{G}_2^\top|, \end{aligned} \quad (44)$$

where (a) follows because  $(\mathbf{X}_1, \mathbf{Z}_2)$  and  $\mathbf{X}_2$  are independent. Combining (43) with (44) yields

$$I(\mathbf{U}; \mathbf{Y}_1) - I(\mathbf{U}; \mathbf{V}) - I(\mathbf{U}; \mathbf{Y}_2 | \mathbf{V}) = \frac{1}{2} \log |\mathbf{I} + \mathbf{G}_1 \mathbf{K}^* \mathbf{G}_1^\top| - \frac{1}{2} \log |\mathbf{I} + \mathbf{G}_2 \mathbf{K}^* \mathbf{G}_2^\top|, \quad (45)$$

which implies that  $(\mathcal{H}_1^K, 0) \in \hat{\mathcal{I}}_K$ . Furthermore, since  $(\mathcal{H}_1^K, 0)$  is on the boundary of  $\hat{\mathcal{O}}_K$  and  $\hat{\mathcal{I}}_K \subseteq \hat{\mathcal{O}}_K$ ,  $(\mathcal{H}_1^K, 0)$  must also be a boundary point of  $\hat{\mathcal{I}}_K$ , and therefore,  $R_1 \leq \mathcal{H}_1^K$  is a supporting hyperplane of  $\hat{\mathcal{I}}_K$ . ■

Based on Lemma 1, to show that the regions coincide, it suffices to show that

$$\max_{(R_1, R_2) \in \hat{\mathcal{O}}_K} \lambda_1 R_1 + \lambda_2 R_2 \leq \max_{(R_1, R_2) \in \hat{\mathcal{I}}_K} \lambda_1 R_1 + \lambda_2 R_2, \quad (46)$$

for  $\lambda_1, \lambda_2 > 0$ . Observe that

$$\begin{aligned}
& \max_{(R_1, R_2) \in \hat{\mathcal{O}}_K} \lambda_1 R_1 + \lambda_2 R_2 \\
& \stackrel{(a)}{\leq} \sup_{\substack{(V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_1 \left[ I(U; \mathbf{Y}_1|V) - I(U; \mathbf{Y}_2|V) \right] + \lambda_2 I(V; \mathbf{Y}_2) \\
& \stackrel{(b)}{=} \sup_{\substack{(V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2|V) \\
& \quad + \lambda_1 \left[ I(\mathbf{X}; \mathbf{Y}_2|V, U) - I(\mathbf{X}; \mathbf{Y}_1|V, U) \right] + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) \\
& \stackrel{(c)}{\leq} \sup_{\substack{V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2|V) + \lim_{\eta \downarrow 1} \lambda_1 S_\eta^Q(\mathbf{X}|V) + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) \\
& \leq \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\substack{V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lim_{\eta \downarrow 1} t_{\lambda, \eta}(\mathbf{X}|V) \\
& \stackrel{(d)}{=} \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} \lim_{\eta \downarrow 1} T_{\lambda, \eta}(\mathbf{X}) \\
& \stackrel{(e)}{=} \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \lim_{\eta \downarrow 1} \hat{V}_{\lambda, \eta}(\mathbf{K}), \tag{47}
\end{aligned}$$

where:

(a) uses (38);

(b) is because  $(V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain;

(c) follows by the definition of  $S_\eta^Q(\mathbf{X}|V)$  and since conditioned on  $V$ ,  $U - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, i.e., it holds that  $P_{\mathbf{Y}_1, \mathbf{Y}_2|V, U, \mathbf{X}} = P_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}}$ . Furthermore, (c) uses the continuity of  $S_\eta^Q(\mathbf{X}|V)$  in  $\eta$  at  $\eta = 1$  (see Property 3 of Proposition 1), which implies that for any  $(V, \mathbf{X})$

$$S_1^Q(\mathbf{X}|V) \triangleq S_{\lim_{\eta \downarrow 1} \eta}^Q(\mathbf{X}|V) = \lim_{\eta \downarrow 1} S_\eta^Q(\mathbf{X}|V);$$

(d) is by the definition of  $T_{\lambda, \eta}^Q(\mathbf{X})$ , the Markov relation  $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ , and because  $T_{\lambda, \eta}^Q(\mathbf{X})$  is continuous in  $\eta$  at  $\eta = 1$  (see Remark 8);

(e) follows by Proposition 3.

By Theorem 4, for every  $\eta > 1$ , there exist independent random variables  $\mathbf{X}_1^* \sim \mathcal{N}(\mathbf{0}, K_1)$ ,  $\mathbf{X}_2^* \sim \mathcal{N}(\mathbf{0}, K_2)$  and  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K - (K_1 + K_2))$ ,  $K_1 + K_2 \preceq K$ , such that  $\hat{V}_\eta^Q(K) = t_{\lambda, \eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*)$  and  $S_\eta^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = s_\eta^Q(\mathbf{X}_1^*)$ . Moreover, setting  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}'$  maximizes  $\lambda_2 I(\mathbf{X}; \mathbf{Y}_2)$  and attains  $\hat{V}_\eta^Q(K)$  simultaneously. In order to conform to the notation in the bounds, let  $V^* = \mathbf{X}'$ . Taking the limit as  $\eta \downarrow 1$ , we have

$$\begin{aligned}
& \max_{(R_1, R_2) \in \hat{\mathcal{O}}_K} \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V^*) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2|V^*) + \lambda_1 \left[ I(\mathbf{X}; \mathbf{Y}_2|V^*, \mathbf{X}_2^*) - I(\mathbf{X}; \mathbf{Y}_2|V^*, \mathbf{X}_2^*) \right] + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) \\
& \leq \lambda_1 \left[ I(\mathbf{X}_2^*; \mathbf{Y}_1|V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2|V^*) \right] + \lambda_2 I(V^*; \mathbf{Y}_2). \tag{48}
\end{aligned}$$

**Proposition 8 (Partial Dirty-Paper Coding (P-DPC))** Let  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + V^*$ , where  $\mathbf{X}_1^*$ ,  $\mathbf{X}_2^*$  and  $V$  are independent Gaussian random vectors with covariances  $\mathbf{K}_1$ ,  $\mathbf{K}_2$  and  $\mathbf{K} - (\mathbf{K}_1 + \mathbf{K}_2)$ , respectively, for some  $0 \preceq \mathbf{K}_1, \mathbf{K}_2 \preceq \mathbf{K}$  with  $\mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{K}$ . Let  $\mathbf{Y}_1 = \mathbf{G}_1 \mathbf{X} + \mathbf{Z}_1$ , where  $\mathbf{Z}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  is independent of  $(\mathbf{X}_1^*, \mathbf{X}_2^*, V^*)$ . Set  $U = \mathbf{X}_2^* + \mathbf{A}V^*$ , where  $\mathbf{A} = \mathbf{K}_2 \tilde{\mathbf{G}}_1^\top [\mathbf{I} + \tilde{\mathbf{G}}_1 \mathbf{K}_2 \tilde{\mathbf{G}}_1^\top]^{-1}$  and  $\tilde{\mathbf{G}}_1 = [\mathbf{I} + \mathbf{G}_1 \mathbf{K}_1 \mathbf{G}_1^\top]^{-\frac{1}{2}} \mathbf{G}_1$ . Then

$$I(\mathbf{X}_2^*; \mathbf{Y}_1 | V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2 | V^*) = I(U^*; \mathbf{Y}_1) - I(U^*; V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2 | V^*). \quad (49)$$

*Proof:* Consider

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{G}_1 \mathbf{X} + \mathbf{Z}_1 \\ &= \mathbf{G}_1 (\mathbf{X}_1^* + \mathbf{X}_2^* + V^*) + \mathbf{Z}_1 \\ &= \mathbf{G}_1 (\mathbf{X}_2^* + V^*) + (\mathbf{G}_1 \mathbf{X}_1^* + \mathbf{Z}_1) \\ &\stackrel{(a)}{=} \mathbf{G}_1 \tilde{\mathbf{X}} + \mathbf{Z}'_1, \end{aligned} \quad (50)$$

where (a) follows by setting  $\tilde{\mathbf{X}} \triangleq \mathbf{X}_2^* + V^*$  and  $\mathbf{Z}'_1 \triangleq \mathbf{G}_1 \mathbf{X}_1^* + \mathbf{Z}_1$ . By the independence of  $\mathbf{X}_1^*$ ,  $\mathbf{X}_2^*$ ,  $V^*$  and  $\mathbf{Z}_1$ , we have that  $\tilde{\mathbf{X}}$  and  $\mathbf{Z}'_1$  are also independent. Moreover,  $\mathbf{Z}'_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I} + \mathbf{G}_1 \mathbf{K}_1 \mathbf{G}_1^\top)$ , where the covariance matrix  $\mathbf{I} + \mathbf{G}_1 \mathbf{K}_1 \mathbf{G}_1^\top$  is diagonalizable (due to its symmetry) and invertible (because it is positive-definite). Denoting  $\Sigma \triangleq \mathbf{I} + \mathbf{G}_1 \mathbf{K}_1 \mathbf{G}_1^\top$ , gives

$$\Sigma = \mathbf{Q} \Lambda \mathbf{Q}^\top, \quad (51)$$

where  $\mathbf{Q}$  is a unitary matrix and  $\Lambda$  is diagonal, and furthermore  $\Sigma^{-\frac{1}{2}} = \mathbf{Q} \Lambda^{-\frac{1}{2}} \mathbf{Q}^\top$ . By defining  $\tilde{\mathbf{Y}}_1 = \Sigma^{-\frac{1}{2}} \mathbf{Y}_1$ , we have

$$\tilde{\mathbf{Y}}_1 = \tilde{\mathbf{G}}_1 \tilde{\mathbf{X}} + \tilde{\mathbf{Z}}_1, \quad (52)$$

where  $\tilde{\mathbf{G}}_1 = \Sigma^{-\frac{1}{2}} \mathbf{G}_1$ ,  $\tilde{\mathbf{Z}}_1 = \Sigma^{-\frac{1}{2}} \mathbf{Z}'_1$  and  $\tilde{\mathbf{Z}}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . Setting  $U^*$  as above and invoking the classic Dirty-Paper Coding Theorem (here we use the formulation from [15, Proposition 12]), we have

$$I(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}_1 | V^*) = I(U^*; \tilde{\mathbf{Y}}_1) - I(U^*; V^*). \quad (53)$$

Furthermore,

$$I(\mathbf{X}_2^*; \mathbf{Y}_2 | V^*) = I(\mathbf{X}_2^* + \mathbf{A}V^*; \mathbf{Y}_2 | V^*) = I(U^*; \mathbf{Y}_2 | V^*). \quad (54)$$

Note that  $\mathbf{Y}_1 \mapsto \Sigma^{-\frac{1}{2}} \mathbf{Y}_1$  is an invertible mapping, and as such, preserves mutual information. We conclude the proof as follows:

$$\begin{aligned} I(\mathbf{X}_2^*; \mathbf{Y}_1 | V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2 | V^*) &\stackrel{(a)}{=} I(\tilde{\mathbf{X}}; \mathbf{Y}_1 | V^*) - I(U^*; \mathbf{Y}_2 | V^*) \\ &\stackrel{(b)}{=} I(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}_1 | V^*) - I(U^*; \mathbf{Y}_2 | V^*) \\ &\stackrel{(c)}{=} I(U^*; \tilde{\mathbf{Y}}_1) - I(U^*; V^*) - I(U^*; \mathbf{Y}_2 | V^*) \end{aligned}$$

$$\stackrel{(d)}{=} I(U^*; \mathbf{Y}_1) - I(U^*; V^*) - I(U^*; \mathbf{Y}_2|V^*), \quad (55)$$

where (a) is because  $\tilde{\mathbf{X}} = \mathbf{X}_2^* + V^*$  and by (54), (b) and (d) follow since  $\mathbf{Y}_1 \mapsto \Sigma^{-\frac{1}{2}}\mathbf{Y}_1$  preserves mutual information, while (c) uses (53). ■

Inserting  $U^*$  as stated in Proposition 8 into the RHS of (48), we obtain

$$\begin{aligned} \max_{(R_1, R_2) \in \hat{\mathcal{O}}_K} \lambda_1 R_1 + \lambda_2 R_2 &\leq \lambda_1 \left[ I(\mathbf{X}_2^*; \mathbf{Y}_1|V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2|V^*) \right] + \lambda_2 I(V^*; \mathbf{Y}_2) \\ &= \lambda_1 \left[ I(U^*; \mathbf{Y}_1) - I(U^*; V^*) - I(U^*; \mathbf{Y}_2|V^*) \right] + \lambda_2 I(V^*; \mathbf{Y}_2) \\ &\stackrel{(a)}{\leq} \max_{(R_1, R_2) \in \hat{\mathcal{I}}_K} \lambda_1 R_1 + \lambda_2 R_2, \end{aligned} \quad (56)$$

where (a) follows since  $(U^*, V^*) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain and  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}$  is satisfied, which implies that the rate pair  $R_1 = I(U^*; \mathbf{Y}_1) - I(U^*; V^*) - I(U^*; \mathbf{Y}_2|V^*)$  and  $R_2 = I(V^*; \mathbf{Y}_2)$  belongs to  $\hat{\mathcal{I}}_K$ . Concluding, we see that  $\hat{\mathcal{I}}_K = \hat{\mathcal{C}}_K = \hat{\mathcal{O}}_K$ , which characterizes the secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages.

Furthermore, equality (and hence the extreme points of  $\hat{\mathcal{C}}_K$ ) is achieved by Gaussian inputs as stated in Proposition 8, thus making the region computable. By evaluating  $\hat{\mathcal{I}}_K$  (or, equivalently  $\hat{\mathcal{O}}_K$ ) with respect to this input distribution, we describe the secrecy-capacity region  $\hat{\mathcal{C}}_K$  as the union of rate pairs  $(R_1, R_2) \in \mathbb{R}_+^2$  satisfying:

$$R_1 \leq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_1\mathbf{K}_1\mathbf{G}_1^\top} \right| - \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2\mathbf{K}_1\mathbf{G}_2^\top} \right| \quad (57a)$$

$$R_2 \leq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2\mathbf{K}\mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{G}_2^\top} \right|, \quad (57b)$$

where the union is over all positive semi-definite matrices  $\mathbf{K}_1, \mathbf{K}_2$ , such that  $\mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{K}$ . We further simplify (57) by noting that the RHS of (57a) is the secrecy-capacity of the MIMO Gaussian WTC as derived in [15, Appendix III], which is maximized by setting  $\mathbf{K}_1 = 0$  (see [6]–[8]). Further note that  $\mathbf{K}_1 = 0$  cannot decrease the RHS of (57b). Thus, by relabeling  $\mathbf{K}_2 \triangleq \mathbf{K}^*$  we establish (6).

## B. Proof of Theorem 2

As in for the case without a common message, the secrecy-capacity region  $\mathcal{C}_K$  is derived by showing that certain outer bound and inner bounds on  $\mathcal{C}_K$  coincide. Denoting by  $\mathcal{C}$  the region of the DM-BC with common, private and confidential messages, we bound it as follows.

**Bound 3 (Outer Bound)** Let  $\mathcal{O}$  be the closure of the union of rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$  satisfying:

$$R_0 \leq \min \{I(W; Y_1), I(W; Y_2)\} \quad (58a)$$

$$R_1 \leq I(U; Y_1|W, V) - I(U; Y_2|W, V) \quad (58b)$$

$$R_0 + R_2 \leq I(V; Y_2|W) + \min \{I(W; Y_1), I(W; Y_2)\} \quad (58c)$$

over all  $(W, V, U) - X - (Y_1, Y_2)$ . Then  $\mathcal{C} \subseteq \mathcal{O}$ .

**Bound 4 (Inner Bound)** Let  $\mathcal{I}$  be the closure of the union of rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$  satisfying:

$$R_0 \leq \min \{I(W; Y_1), I(W; Y_2)\} \quad (59a)$$

$$R_1 \leq I(U; Y_1|W) - I(U; V|W) - I(U; Y_2|W, V) \quad (59b)$$

$$R_2 \leq I(V; Y_2|W) \quad (59c)$$

over all  $(W, V, U) - X - (Y_1, Y_2)$ . Then  $\mathcal{I} \subseteq \mathcal{C}$ .

The proofs of Bounds 3 and 4 are relegated to Appendix A. Denoting by  $\mathcal{O}_K$  and  $\mathcal{I}_K$  the adaptations of Bounds 3 and 4 to the case of a MIMO Gaussian BC with a covariance input constraint  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K$ , we have  $\mathcal{I}_K \subseteq \mathcal{C}_K \subseteq \mathcal{O}_K$ .

Next, we use the factorization of concave envelopes method to show that the opposite inclusion, i.e.,  $\mathcal{O}_K \subseteq \mathcal{I}_K$ , also holds. Given the supporting hyperplanes characterization of bounded and closed convex sets, using a similar reasoning as in Section V-A (see Lemma 1), it suffices to characterize  $\max_{(R_0, R_1, R_2) \in \mathcal{C}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$ , for  $\lambda_j > 0$ ,  $j = 1, 2, 3$ . Further note that it suffices to restrict the discussion to the case  $\lambda_0 > \lambda_2$ . This is due to the following observation: If a rate triple  $(R_0, R_1, R_2)$  is in  $\mathcal{C}_K$  then so does the triple  $(0, R_1, R_2 + R_0)$ , since one may always treat the common message as part of the (non-confidential) private message to receiver 2. Assuming  $\lambda_0 \leq \lambda_2$ , we have

$$\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq 0 \cdot R_0 + \lambda_1 R_1 + \lambda_2 (R_0 + R_2), \quad (60)$$

and therefore,

$$\max_{(R_0, R_1, R_2) \in \mathcal{C}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 = \max_{(0, R_1, R_2) \in \mathcal{C}_K} \lambda_1 R_1 + \lambda_2 R_2 = \max_{(R_1, R_2) \in \hat{\mathcal{C}}_K} \lambda_1 R_1 + \lambda_2 R_2, \quad (61)$$

where  $\hat{\mathcal{C}}_K$  is the secrecy-capacity region without a common message that was characterized in Section III-A.

Hence, it suffices to show that for all  $\lambda_j > 0$ ,  $j = 1, 2, 3$ , with  $\lambda_0 > \lambda_2$ , we have

$$\max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq \max_{(R_0, R_1, R_2) \in \mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2. \quad (62)$$

Now, for any  $\alpha \in [0, 1]$  set  $\bar{\alpha} = 1 - \alpha$ , and consider the following.

$$\begin{aligned} & \max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \stackrel{(a)}{\leq} \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_0 \left[ \alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1|W, V) - I(U; \mathbf{Y}_2|W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2|W) \end{aligned}$$

$$\begin{aligned}
& \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2|W) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1|W) \\
\stackrel{(b)}{=} & \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K}} \quad + \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|W, V) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|W, V) \\
& \quad + \lambda_1 I(\mathbf{X}; \mathbf{Y}_2|W, V, U) - \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|W, V, U) \\
\stackrel{(c)}{\leq} & \sup_{\substack{(W,V)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K}} \quad \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2|W) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1|W) \\
& \quad + \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|W, V) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|W, V) + \lim_{\eta \downarrow 1} \lambda_1 S_\eta^Q(\mathbf{X}|W, V) \\
\stackrel{(d)}{\leq} & \sup_{\substack{W-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K}} \quad \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) \\
& \quad + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2|W) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1|W) + \lim_{\eta \downarrow 1} T_{\lambda,\eta}^Q(\mathbf{X}|W) \\
\leq & \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K} \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\substack{W-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K}} \lim_{\eta \downarrow 1} f_{\lambda_0,\alpha,\eta}^Q(\mathbf{X}|W) \\
\stackrel{(e)}{\leq} & \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K} \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K} \lim_{\eta \downarrow 1} F_{\lambda_0,\alpha,\eta}^Q(\mathbf{X}) \\
\stackrel{(f)}{\leq} & \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \leq K} \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + \lim_{\eta \downarrow 1} \tilde{V}_{\lambda_0,\alpha,\eta}^Q(K), \tag{63}
\end{aligned}$$

where:

(a) is by (58);

(b) follows because  $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain;

(c) is by the definition of  $S_\eta^Q(\mathbf{X}|W, V)$  since conditioned on  $(W, V)$ ,  $U - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, and because  $S_\eta^Q(\mathbf{X}|V, W)$  is continuous in  $\eta$  at  $\eta = 1$  (Property 3 of Proposition 1);

(d) follows by the definition of  $T_{\lambda,\eta}^Q(\mathbf{X}|W)$  since conditioned on  $V$ ,  $W - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain. Furthermore, the continuity of  $T_{\lambda,\eta}^Q(\mathbf{X}|W)$  in  $\eta$  at  $\eta = 1$  (see Remark 8) is also exploited;

(e) is by the definition of  $F_{\lambda_0,\alpha,\eta}^Q(\mathbf{X})$  (while noting that  $W - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain), and because  $F_{\lambda,\eta}^Q(\mathbf{X})$  is continuous at  $\eta = 1$  (Remark 9);

(f) makes use of the continuity argument from Remark 10.

Recall that for any  $\eta > 1$ ,  $\lambda_j > 0$ , for  $j = 0, 1, 2$ , and  $\lambda_0 > \lambda_2$ , Corollary 3 implies that there exist independent random variables  $\mathbf{X}_j^* \sim \mathcal{N}(\mathbf{0}, K_j)$ ,  $j = 1, 2, 3$ , and  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K - (K_1 + K_2 + K_3))$ , such that (37) is satisfied. Furthermore, setting  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^* + \mathbf{X}'$  not only attains  $\tilde{V}_\eta^Q(K)$ , but it also simultaneously maximizes  $\alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1)$  and  $\bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2)$ . Relabeling  $W^* = \mathbf{X}'$  and  $V^* = \mathbf{X}_3^*$  while taking the limit as  $\eta \downarrow 1$ , we have

$$\begin{aligned}
& \max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \bar{\alpha}\lambda_0 I(\mathbf{X}; \mathbf{Y}_2) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2|W^*) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1|W^*) \\
& \quad + \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|W^*, V^*) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|W^*, V^*) + \lambda_1 I(\mathbf{X}; \mathbf{Y}_2|W^*, V^*, \mathbf{X}_2^*) - \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|W^*, V^*, \mathbf{X}_2^*) \\
& \leq \lambda_0 \left[ \alpha I(W^*; \mathbf{Y}_1) + \bar{\alpha} I(W^*; \mathbf{Y}_2) \right] + \lambda_1 \left[ I(\mathbf{X}_2^*; \mathbf{Y}_1|W^*, V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2|W^*, V^*) \right] + \lambda_2 I(V^*; \mathbf{Y}_2|W^*). \tag{64}
\end{aligned}$$

Using Proposition 8, we set  $U = \mathbf{X}_2^* + \tilde{\mathbf{A}}V^*$  as before and obtain

$$I(\mathbf{X}_2^*; \mathbf{Y}_1 | W^*, V^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2 | W^*, V^*) = I(U^*; \mathbf{Y}_1 | W^*) - I(U^*; V^* | W^*) - I(\mathbf{X}_2^*; \mathbf{Y}_2 | W^*, V^*). \quad (65)$$

Inserting (65) into (64), yields

$$\begin{aligned} & \max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \lambda_0 \left[ \alpha I(W^*; \mathbf{Y}_1) + \bar{\alpha} I(W^*; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U^*; \mathbf{Y}_1 | W^*) - I(U^*; V^* | W^*) - I(U^*; \mathbf{Y}_2 | W^*, V^*) \right] + \lambda_2 I(V^*; \mathbf{Y}_2 | W^*) \\ & \leq \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_0 \left[ \alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W). \end{aligned} \quad (66)$$

Since (66) holds for all  $\alpha \in [0, 1]$ , we have

$$\begin{aligned} & \max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \min_{\alpha \in [0, 1]} \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_0 \left[ \alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W). \end{aligned} \quad (67)$$

Having (67), the desired equality  $\mathcal{O}_K = \mathcal{I}_K$  is a consequence of the following Proposition <sup>3</sup>.

**Proposition 9 (Max-Min Interchanging)** *The following max-min interchanging holds*

$$\begin{aligned} & \min_{\alpha \in [0, 1]} \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_0 \left[ \alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W) \\ & = \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \min_{\alpha \in [0, 1]} \lambda_0 \left[ \alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \right] \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W) \\ & = \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} \lambda_0 \cdot \min \{ I(W; \mathbf{Y}_1), I(W; \mathbf{Y}_2) \} \\ & \quad + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W). \end{aligned} \quad (68)$$

The proof of Proposition 9 is given in Appendix C.

Now, noting that  $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain and  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K$ , we see that the triple

$$\begin{aligned} R_0 &= \min \{ I(W; \mathbf{Y}_1), I(W; \mathbf{Y}_2) \} \\ R_1 &= I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \end{aligned}$$

<sup>3</sup>Proposition 9 bears strong resemblance to Proposition 13 from [15] that was essentially established in [22]

$$R_2 = I(V; \mathbf{Y}_2|W)$$

is inside the inner bound  $\mathcal{I}_K$ . Hence

$$\begin{aligned} & \max_{(R_0, R_1, R_2) \in \mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \sup_{\substack{(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}}} \lambda_0 [\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)] \\ & \quad + \lambda_1 [I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)] + \lambda_2 I(V; \mathbf{Y}_2|W) \\ & \leq \max_{(R_0, R_1, R_2) \in \mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2, \end{aligned} \quad (69)$$

which implies that  $\hat{\mathcal{I}}_K = \hat{\mathcal{C}}_K = \hat{\mathcal{O}}_K$  and characterizes the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages.

To obtain the description of  $\mathcal{C}_K$  stated in (10), note that when  $\lambda_0 > \lambda_2$ , equality (and hence the extreme points of  $\mathcal{C}_K$ ) is achieved by setting

$$\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + V^* + W^*, \quad (70)$$

where  $\mathbf{X}_j^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_j)$ ,  $j = 1, 2$ ,  $V^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_3)$  and  $W^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K} - (\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3))$  are independent of each other, and  $U = \mathbf{X}_2^* + \tilde{\mathbf{A}}V^*$ , where  $\tilde{\mathbf{A}}$  is the P-DPC matrix from Proposition 8. For the case when  $\lambda_0 \leq \lambda_2$ , (60)-(61) imply that the boundary-achieving input distribution corresponds to the one that achieves the secrecy-capacity region when there is no common message (see Section V-A). Setting  $\mathbf{K}_3 = \mathbf{K} - (\mathbf{K}_1 + \mathbf{K}_2)$  recovers the optimal input distribution for the case without common message.

By evaluating  $\mathcal{I}_K$  (or, equivalently  $\mathcal{O}_K$ ) with respect to (70), we characterize the secrecy-capacity region  $\mathcal{C}_K$  as the union of rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$  satisfying:

$$R_0 \leq \min \left\{ \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1 \mathbf{K} \mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_1 (\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3) \mathbf{G}_1^\top} \right|, \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1 \mathbf{K} \mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_1 (\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3) \mathbf{G}_1^\top} \right| \right\} \quad (71a)$$

$$R_1 \leq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_1 (\mathbf{K}_1 + \mathbf{K}_2) \mathbf{G}_1^\top}{\mathbf{I} + \mathbf{G}_1 \mathbf{K}_1 \mathbf{G}_1^\top} \right| - \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2 (\mathbf{K}_1 + \mathbf{K}_2) \mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2 \mathbf{K}_1 \mathbf{G}_2^\top} \right| \quad (71b)$$

$$R_2 \leq \frac{1}{2} \log \left| \frac{\mathbf{I} + \mathbf{G}_2 (\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3) \mathbf{G}_2^\top}{\mathbf{I} + \mathbf{G}_2 (\mathbf{K}_1 + \mathbf{K}_2) \mathbf{G}_2^\top} \right|, \quad (71c)$$

where the union is over all positive semi-definite matrices  $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3$ , such that  $\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3 \preceq \mathbf{K}$ .

The region from (71) is further simplified using reasoning similar to this from Section V-A. First, (71b) indicates that the signal to User 1 is a sum of two independent zero mean Gaussian random vector with covariance  $\mathbf{K}_1$  and  $\mathbf{K}_2$ . The signal that corresponds to  $\mathbf{K}_2$  carries the confidential message  $M_1$ , while the  $\mathbf{K}_1$  signal is an artificial noise sent (on purpose) to confuse User 2 (which serves as an eavesdropper of  $M_1$ ). The lack of structure in the artificial noise adds to the noise floor at both receivers (see also [12, Remark 4]). However, since the RHS of (71b) is the secrecy-capacity of the MIMO Gaussian WTC, it is maximized by setting  $\mathbf{K}_1 = \mathbf{0}$ . Furthermore, since in both (71a) and (71c)  $\mathbf{K}_1$  serves as noise (i.e., it is not used to encode any of the messages), setting  $\mathbf{K}_1 = \mathbf{0}$  achieves optimality. This is since  $\mathbf{K}_1 = \mathbf{0}$  corresponds to revealing the  $\mathbf{K}_1$  signal to both receivers, which can only increase



the transmission rates. Taking  $K_1 = 0$  and recasting  $K_3$  as  $K_1$ , recovers (10).

## VI. PROOFS OF UPPER CONCAVE ENVELOPES PROPERTIES

### A. Proof of Proposition 1

Property 1 follows by Jensen's inequality since  $S_\eta^Q$  is concave in  $P_{\mathbf{X}}$ , while for Property 2 we use the fact that  $P_{\mathbf{X}|W,V} = P_{\mathbf{X}|V}$ . To prove Property 3, fix  $P_{\mathbf{X}}$  and let  $\eta_1, \eta_2 \in (0, 2)$ ,  $\alpha \in [0, 1]$  and  $\bar{\alpha} = 1 - \alpha$ . Observe that

$$\begin{aligned} S_{\alpha\eta_1 + \bar{\alpha}\eta_2}^Q(\mathbf{X}) &= \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - (\alpha\eta_1 + \bar{\alpha}\eta_2)I(\mathbf{X}; \mathbf{Y}_1|V) \\ &\leq \alpha \cdot \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - \eta_1 I(\mathbf{X}; \mathbf{Y}_1|V) + \bar{\alpha} \cdot \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - \eta_2 I(\mathbf{X}; \mathbf{Y}_1|V) \\ &= \alpha S_{\eta_1}^Q(\mathbf{X}) + \bar{\alpha} S_{\eta_2}^Q(\mathbf{X}). \end{aligned}$$

Clearly,  $S_\eta^Q(\mathbf{X})$  is also bounded for every  $\eta \in (0, 2)$  and by invoking Proposition 17 from [20, Chapter 5], we have that  $S_\eta^Q(\mathbf{X})$  is continuous inside every closed subinterval of  $(0, 2)$ , and in particular, at  $\eta = 1$ .

### B. Proof of Proposition 4

Let  $V - (\mathbf{X}_1, \mathbf{X}_2) - (\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Y}_{21}, \mathbf{Y}_{22})$  form a Markov chain. We have,

$$\begin{aligned} t_{\lambda, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V) &= \lambda_1 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}|V) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}|V) + \lambda_1 S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\ &= \lambda_1 \left[ I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}|V) + I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) \right] \\ &\quad - (\lambda_1 + \lambda_2) \left[ I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{22}|V) + I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) \right] + \lambda_1 S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\ &\stackrel{(a)}{=} \lambda_1 \left[ I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) + I(\mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{11}) \right] \\ &\quad - (\lambda_1 + \lambda_2) \left[ I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) + I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \right] + \lambda_1 S_\eta^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\ &\stackrel{(b)}{\leq} \lambda_1 I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) + \lambda_1 S_\eta^{Q_1}(\mathbf{X}_1|V, \mathbf{Y}_{22}) \\ &\quad + \lambda_1 I(\mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) - (\lambda_1 + \lambda_2) I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) + \lambda_1 S_\eta^{Q_2}(\mathbf{X}_2|V, \mathbf{Y}_{11}) - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\ &\stackrel{(c)}{\leq} T_{\lambda, \eta}^{Q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + T_{\lambda, \eta}^{Q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\ &\stackrel{(d)}{\leq} T_{\lambda, \eta}^{Q_1}(\mathbf{X}_1) + T_{\lambda, \eta}^{Q_2}(\mathbf{X}_2) - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\ &\leq T_{\lambda, \eta}^{Q_1}(\mathbf{X}_1) + T_{\lambda, \eta}^{Q_2}(\mathbf{X}_2), \end{aligned} \tag{72}$$

where:

(a) is since given  $V$  we have the Markov chain  $(\mathbf{Y}_{11}, \mathbf{Y}_{21}) - \mathbf{X}_1 - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$ ;

(b) follows from Proposition 2 by the definition of  $S_\eta^Q(\cdot|\cdot)$ ;

(c) is because  $(V, \mathbf{Y}_{22}) - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$  and  $(V, \mathbf{Y}_{11}) - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$  form Markov chains;

(d) follows by Remark 8 due to the Markov chains  $\mathbf{Y}_{22} - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$  and  $\mathbf{Y}_{11} - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$ .

Now for  $(V^*, \mathbf{X}_1^*, \mathbf{X}_2^*)$ , an end-to-end equality holds in (72). In particular, this implies that  $I(\mathbf{Y}_{11}^*; \mathbf{Y}_{22}^* | V^*) = 0$ , i.e., that  $\mathbf{Y}_{11}^* - V^* - \mathbf{Y}_{22}^*$  forms a Markov chain. By Proposition 2 in [15], we have that  $\mathbf{X}_1^* - V^* - \mathbf{X}_2^*$ , which further implies the Markov chain

$$(\mathbf{Y}_{11}^*, \mathbf{Y}_{21}^*) - \mathbf{X}_1^* - V^* - \mathbf{X}_2^* - (\mathbf{Y}_{12}^*, \mathbf{Y}_{22}^*). \quad (73)$$

The end-to-end equality in (72) also gives

$$\begin{aligned} T_{\lambda, \eta}^{Q_1}(\mathbf{X}_1^*) &= \lambda_1 I(\mathbf{X}_1^*; \mathbf{Y}_{11}^* | V^*, \mathbf{Y}_{22}^*) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1^*; \mathbf{Y}_{21}^* | V^*, \mathbf{Y}_{22}^*) + \lambda_1 S_{\eta}^{Q_1}(\mathbf{X}_1^* | V^*, \mathbf{Y}_{22}^*) \\ &\stackrel{(a)}{=} \lambda_1 I(\mathbf{X}_1^*; \mathbf{Y}_{11}^* | V^*) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1^*; \mathbf{Y}_{21}^* | V^*) + \lambda_1 S_{\eta}^{Q_1}(\mathbf{X}_1^* | V^*), \end{aligned} \quad (74)$$

where (a) follows because (73) implies that the chain  $\mathbf{Y}_{22}^* - V^* - \mathbf{X}_1^* - (\mathbf{Y}_{11}^*, \mathbf{Y}_{21}^*)$  is Markov. Similarly, it can be shown that  $T_{\lambda, \eta}^{Q_2}(\mathbf{X}_2^*) = t_{\lambda, \eta}^{Q_2}(\mathbf{X}_2^* | V^*)$ .

### C. Proof of Proposition 5

Let  $V - (\mathbf{X}_1, \mathbf{X}_2) - (\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Y}_{21}, \mathbf{Y}_{22})$  be a Markov chain. We have,

$$\begin{aligned} f_{\lambda_0, \alpha, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2 | V) &= (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) - \alpha \lambda_0 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) + T_{\lambda, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2 | V) \\ &\stackrel{(a)}{=} (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}_1; \mathbf{Y}_{21} | V, \mathbf{Y}_{22}) - \alpha \lambda_0 I(\mathbf{X}_1; \mathbf{Y}_{11} | V, \mathbf{Y}_{22}) - \alpha \lambda_0 I(\mathbf{Y}_{11}; \mathbf{Y}_{22} | V) \\ &\quad + (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}_2; \mathbf{Y}_{22} | V, \mathbf{Y}_{11}) - \alpha \lambda_0 I(\mathbf{X}_2; \mathbf{Y}_{21} | V, \mathbf{Y}_{11}) + (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{Y}_{11}; \mathbf{Y}_{22} | V) + T_{\lambda, \eta}^{Q_1 \times Q_2}(\mathbf{X}_1, \mathbf{X}_2 | V) \\ &\stackrel{(b)}{\leq} (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}_1; \mathbf{Y}_{21} | V, \mathbf{Y}_{22}) - \alpha \lambda_0 I(\mathbf{X}_1; \mathbf{Y}_{11} | V, \mathbf{Y}_{22}) + T_{\lambda, \eta}^{Q_1}(\mathbf{X}_1 | V, \mathbf{Y}_{22}) \\ &\quad + (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}_2; \mathbf{Y}_{22} | V, \mathbf{Y}_{11}) - \alpha \lambda_0 I(\mathbf{X}_2; \mathbf{Y}_{21} | V, \mathbf{Y}_{11}) + T_{\lambda, \eta}^{Q_2}(\mathbf{X}_2 | V, \mathbf{Y}_{11}) - (\lambda_0 - \lambda_2) I(\mathbf{Y}_{11}; \mathbf{Y}_{22} | V) \\ &\stackrel{(c)}{\leq} F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1 | \mathbf{Y}_{22}) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2 | \mathbf{Y}_{11}) - (\lambda_0 - \lambda_2) I(\mathbf{Y}_{11}; \mathbf{Y}_{22} | V) \\ &\stackrel{(d)}{\leq} F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2) - (\lambda_0 - \lambda_2) I(\mathbf{Y}_{11}; \mathbf{Y}_{22} | V) \\ &\leq F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1) + F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2), \end{aligned} \quad (75)$$

where:

(a) is similar to step (a) in the proof of Proposition 4;

(b) uses Proposition 4 and the definition of  $T_{\lambda, \eta}^Q(\cdot | \cdot)$ ;

(c) is because  $(V, \mathbf{Y}_{22}) - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$  and  $(V, \mathbf{Y}_{11}) - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$  form Markov chains;

(d) follows by Remark 9 due to the Markov chains  $\mathbf{Y}_{22} - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$  and  $\mathbf{Y}_{11} - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$ .

For  $(V^*, \mathbf{X}_1^*, \mathbf{X}_2^*)$  that satisfy (35), an end-to-end equality holds in (75), implying that  $I(\mathbf{Y}_{11}^*; \mathbf{Y}_{22}^* | V^*) = 0$ . Invoking Proposition 2 from [15], we deduce that  $\mathbf{X}_1^* - V^* - \mathbf{X}_2^*$  forms a Markov chain, which further implies the

Markov chain

$$(\mathbf{Y}_{11}^*, \mathbf{Y}_{21}^*) - \mathbf{X}_1^* - V^* - \mathbf{X}_2^* - (\mathbf{Y}_{12}^*, \mathbf{Y}_{22}^*). \quad (76)$$

By the end-to-end equality in (72), we also have

$$\begin{aligned} F_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1^*) &= (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1^*; \mathbf{Y}_{21}^*|V^*, \mathbf{Y}_{22}^*) - \alpha\lambda_0 I(\mathbf{X}_1^*; \mathbf{Y}_{11}^*|V^*, \mathbf{Y}_{22}^*) + T_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1^*|V^*, \mathbf{Y}_{22}^*) \\ &\stackrel{(a)}{=} (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1^*; \mathbf{Y}_{21}^*|V^*) - \alpha\lambda_0 I(\mathbf{X}_1^*; \mathbf{Y}_{11}^*|V^*) + T_{\lambda_0, \alpha, \eta}^{Q_1}(\mathbf{X}_1^*|V^*), \end{aligned} \quad (77)$$

where (a) uses (76), which implies that  $\mathbf{Y}_{22}^* - V^* - \mathbf{X}_1^* - (\mathbf{Y}_{11}^*, \mathbf{Y}_{21}^*)$  forms a Markov chain. Similarly, it can be shown that  $F_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2^*) = f_{\lambda_0, \alpha, \eta}^{Q_2}(\mathbf{X}_2^*|V^*)$ .

#### D. Proof of Proposition 7

Denote  $Q_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}} \triangleq Q$  and consider the two-letter BC  $Q(\mathbf{y}_{11}, \mathbf{y}_{21}|\mathbf{x}_1) \times Q(\mathbf{y}_{12}, \mathbf{y}_{22}|\mathbf{x}_2)$ . We have

$$\begin{aligned} \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) &\stackrel{(a)}{=} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_1|V_1) + f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_2|V_2) \\ &\stackrel{(b)}{=} f_{\lambda_0, \alpha, \eta}^{Q \times Q}(\mathbf{X}_1, \mathbf{X}_2|V_1, V_2) \\ &\stackrel{(c)}{=} f_{\lambda_0, \alpha, \eta}^{Q \times Q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2}|\tilde{V}) \\ &\stackrel{(d)}{\leq} F_{\lambda_0, \alpha, \eta}^{Q \times Q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2}) \\ &\stackrel{(e)}{\leq} F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_1}) + F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_2}) \\ &\stackrel{(f)}{\leq} \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) + \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = 2\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}), \end{aligned} \quad (78)$$

where:

- (a) is because  $P_{V, \mathbf{X}}^*$  achieves  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ ;
- (b) uses the independence of  $(V_1, \mathbf{X}_1)$  and  $(V_2, \mathbf{X}_2)$ ;
- (c) is a consequence of [15, Proposition 1] (namely, the invariance with respect to rotation of the mutual information between the input and output of an additive Gaussian channel);
- (d) follows by the definition of the double-nested UCE;
- (e) uses Proposition 5;
- (f) is the definition of  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ , while noting that the independence of  $\mathbf{X}_{v_1}$  and  $\mathbf{X}_{v_2}$ , for every  $v_1 \neq v_2 \in \mathcal{V}$ , implies that

$$\begin{aligned} \mathbb{E}[\mathbf{X}_{\theta_1} \mathbf{X}_{\theta_1}^\top] &= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{\sqrt{2}}(\mathbf{X}_{V_1} + \mathbf{X}_{V_2}) \cdot \frac{1}{\sqrt{2}}(\mathbf{X}_{V_1} + \mathbf{X}_{V_2})^\top \middle| \tilde{V}\right]\right] \\ &= \mathbb{E}\left[\frac{1}{2}\mathbb{E}[\mathbf{X}_{V_1} \mathbf{X}_{V_1}^\top | \tilde{V}] + \frac{1}{2}\mathbb{E}[\mathbf{X}_{V_2} \mathbf{X}_{V_2}^\top | \tilde{V}]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{\sqrt{2}}(\mathbf{X}_{V_1} - \mathbf{X}_{V_2}) \cdot \frac{1}{\sqrt{2}}(\mathbf{X}_{V_1} - \mathbf{X}_{V_2})^\top \middle| \tilde{V}\right]\right] \end{aligned}$$

$$= \mathbb{E}[\mathbf{X}_{\theta_2} \mathbf{X}_{\theta_2}^\top], \quad (79)$$

and

$$\mathbb{E}[\mathbf{X}_{\theta_2} \mathbf{X}_{\theta_2}^\top] = \sum_{v_1, v_2} P_V^*(v_1) P_V^*(v_2) \left( \frac{1}{2} \mathbf{K}_{v_1} + \frac{1}{2} \mathbf{K}_{v_2} \right) = \sum_v P_V^*(v) \mathbf{K}_v = \mathbb{E}[\mathbb{E}[\mathbf{X} \mathbf{X}^\top | V]] = \mathbb{E}[\mathbf{X} \mathbf{X}^\top] \preceq \mathbf{K}. \quad (80)$$

In (80) we denote  $\mathbf{K}_v \triangleq \mathbb{E}[\mathbf{X}_v \mathbf{X}_v^\top]$ .

Since the extremes of the chain of inequalities in (78) match, all inequalities are, in fact, equalities. Equality in step (d) implies that  $P_{\tilde{V}|\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2}}$  achieves  $F_{\lambda_0, \alpha, \eta}^{Q \times Q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2})$ . Furthermore, by Proposition 5, since (d) and (e) are equalities we have that  $\mathbf{X}_{\theta_1} - \tilde{V} - \mathbf{X}_{\theta_2}$ , and that  $P_{\tilde{V}|\mathbf{X}_{\theta_1}}$  and  $P_{\tilde{V}|\mathbf{X}_{\theta_2}}$  achieve  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_1})$  and  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_2})$ , respectively. Finally, equality in (f) means that  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_j} | \tilde{V}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{\theta_j} | \tilde{V})$ , for  $j = 1, 2$ .

### E. Proof of Theorem 5

As a consequence of Proposition 7, for any fixed  $(v_1, v_2) \in \mathcal{V}^2$ ,  $\mathbf{X}_{v_1} + \mathbf{X}_{v_2}$  and  $\mathbf{X}_{v_1} - \mathbf{X}_{v_2}$  are independent. Combined with  $\mathbf{X}_{v_1}$  and  $\mathbf{X}_{v_2}$  being independent zero mean random variables, Corollary 3 in Appendix I-A of [15] implies that  $\mathbf{X}_{v_1}$  and  $\mathbf{X}_{v_2}$  are Gaussian random vectors with the same covariance matrix. Since the pair  $(v_1, v_2) \in \mathcal{V}^2$  is arbitrary, we see that all Gaussian vectors  $\{\mathbf{X}_v\}_v$  have the same covariance matrix. Denoting this matrix by  $\mathbf{K}^*$ , it clearly satisfies  $\mathbf{K}^* \preceq \mathbf{K}$ . Letting  $\mathbf{X}^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^*)$ , we have

$$\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) \stackrel{(a)}{=} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X} | V) = \sum_{i=1}^m P_V^*(v_i) f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}_{v_i}) = \sum_{i=1}^m P_V^*(v_i) f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*), \quad (81)$$

where (a) follows since  $(V, \mathbf{X}) \sim P_{V, \mathbf{X}}^*$  attain  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$  (Proposition 6), (b) follows by the definition of  $\mathbf{X}_v$  in the statement of Proposition 7, while (c) follows since  $\mathbf{X}^*$  and  $\mathbf{X}_{v_i}$  are identically distributed, for every  $i \in [1 : m]$ .

To account for the uniqueness of the zero-mean maximizer we first show that if a zero mean random vector  $\mathbf{X}$  is a maximizer, i.e.,  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X})$ , it must be Gaussian. Let  $\mathbf{X}_1$  and  $\mathbf{X}_2$  be two i.i.d. copies of  $\mathbf{X}$ . Applying Proposition 7 while taking  $V$  to be a constant, we obtain that  $\mathbf{X}_1 + \mathbf{X}_2$  and  $\mathbf{X}_1 - \mathbf{X}_2$  are also independent. Hence, by [15, Corollary 3],  $\mathbf{X}$  is Gaussian.

Next, suppose that  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$  has two independent Gaussian maximizers denoted by  $\mathbf{A}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_1)$  and  $\mathbf{A}_2 \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_2)$ , such that  $\mathbf{K}_1, \mathbf{K}_2 \preceq \mathbf{K}$  and  $\mathbf{K}_1 \neq \mathbf{K}_2$ . Let  $(V, \mathbf{X})$  be a pair of random variables, such that  $V \sim \text{Ber}(\frac{1}{2})$  on  $\mathcal{V} = \{1, 2\}$ ,  $\mathbf{X} | \{V = 1\} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_1)$  and  $\mathbf{X} | \{V = 2\} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_2)$ . Note that  $(V, \mathbf{X})$  also attains  $\tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ . Taking  $v_1 = 1$  and  $v_2 = 2$ , Proposition 7 implies that  $\mathbf{A}_1 + \mathbf{A}_2$  and  $\mathbf{A}_1 - \mathbf{A}_2$  are independent, which contradict Corollary 3 from [15] as  $\mathbf{K}_1 \neq \mathbf{K}_2$ .

### F. Proof of Corollary 3

By Theorem 5, there is an  $\mathbf{X}^* \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^*)$ , such that  $\mathbf{K}^* \preceq \mathbf{K}$  and  $f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*) = \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ . Let  $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathbf{K} - \mathbf{K}^*)$  be independent of  $\mathbf{X}^*$  and set  $\mathbf{X} = \mathbf{X}^* + \mathbf{X}'$ . Thus  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$  and by definition we have  $F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) \leq \tilde{V}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K})$ .

On the other hand,

$$F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) = \sup_{\substack{P_{V|\mathbf{X}}: \\ V-\mathbf{X}-(\mathbf{Y}_1, \mathbf{Y}_2)}} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|V) \stackrel{(a)}{\geq} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}|\mathbf{X}') \stackrel{(b)}{=} f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*) = \tilde{V}_{\lambda_0, \alpha, \eta}^Q(K), \quad (82)$$

where (a) follows since  $\mathbf{X}' - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  forms a Markov chain, while (b) follows because  $\mathbf{X}|\{\mathbf{X}' = \mathbf{x}'\} \sim \mathbf{X}^* + \mathbf{x}'$ . Thus,

$$\tilde{V}_{\lambda_0, \alpha, \eta}^Q(K) = F_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^*) = (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}^*; \mathbf{Y}_2) - \alpha\lambda_0 I(\mathbf{X}^*; \mathbf{Y}_1) + T_{\lambda, \eta}^Q(\mathbf{X}^*). \quad (83)$$

By Theorem 4, one can decompose  $\mathbf{X}^*$  into independent  $\mathbf{X}_1^* \sim \mathcal{N}(\mathbf{0}, K_1)$  and  $\mathbf{X}_2^* \sim \mathcal{N}(\mathbf{0}, K_2)$ , with  $K_1 + K_2 \preceq K^*$ , such that

$$T_{\lambda, \eta}^Q(\mathbf{X}^*) = t_{\lambda, \eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = \hat{V}_{\lambda, \eta}^Q(K^*), \quad (84)$$

and

$$S_{\eta}^Q(\mathbf{X}_1^* + \mathbf{X}_2^*) = s_{\eta}^Q(\mathbf{X}_1^*) = V_{\eta}^Q(K_1 + K_2). \quad (85)$$

The proof of existence is concluded by setting  $\mathbf{X}_3^* \sim \mathcal{N}(\mathbf{0}, K_3)$ , where  $K_3 = K^* - (K_1 + K_2)$  and noting that  $\mathbf{X} = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^* + \mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K)$  and  $\mathbf{X}^* = \mathbf{X}_1^* + \mathbf{X}_2^* + \mathbf{X}_3^* \sim \mathcal{N}(\mathbf{0}, K^*)$ , which implies that (37) holds. The uniqueness of the decomposition (i.e., of the covariance matrices  $K_1$ ,  $K_2$  and  $K_3$ ) is a direct consequence of Theorems 3, 4 and 5.

## VII. SUMMARY AND CONCLUDING REMARKS

The two-user MIMO Gaussian BC with common, private and confidential messages was studied. The private message to Receiver 1 is confidential and kept secret from Receiver 2. The secrecy-capacity region of the setting without a common message was characterized first and Gaussian inputs were shown to achieve optimality. The proof relied on establishing an equivalence between certain inner and outer bounds using factorization of UCEs [15] and a variant of DPC [17]. Our results showed that using DPC to cancel out the signal of the non-confidential message at Receiver 1 exhausts the entire region, making DPC against the signal of the confidential message unnecessary.

This secrecy-capacity region without a common message was then used to characterize a portion of the region with a common message. The rest of the region was found using double-nested UCEs. The secrecy-capacity region of the case without a common message was visualized using a numerical example. To make the region (efficiently) computable, matrix decomposition properties from [18] were leveraged. The region was shown to be strictly larger than the secrecy-capacity region of the MIMO Gaussian BC with confidential messages (where each private message is kept secret from the opposite user).

## APPENDIX A

## DERIVATION OF INNER AND OUTER BOUNDS

## A. Outer Bounds 1 and 3

We first establish Bound 3 as an outer bound on the secrecy-capacity region of the setting with a common message, and then use it to establish Bound 1 as an outer bound on the region without a common message.

The result of [21, Theorem 2] characterizes an outer bound  $\mathcal{R}_O(L_1, L_2)$  on the  $(L_1, L_2)$ -leakage-capacity region of a DM-BC with common and private messages, for some allowed leakage thresholds  $(L_1, L_2) \in \mathbb{R}_+^2$ . Setting  $L_1 = 0$  and letting  $L_2 \rightarrow \infty$  in  $\mathcal{R}_O(L_1, L_2)$  (which corresponds to  $M_1$  being confidential and  $M_2$  not being subject to any secrecy requirements), we have that the closure of the union of rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$  satisfying:

$$R_0 \leq \min \{I(W; Y_1), I(W; Y_2)\} \quad (86a)$$

$$R_1 \leq I(U; Y_1|W, V) - I(U; Y_2|W, V) \quad (86b)$$

$$R_1 \leq I(U; Y_1|W) - I(U; Y_2|W) \quad (86c)$$

$$R_0 + R_2 \leq I(V; Y_2|W) + \min \{I(W; Y_1), I(W; Y_2)\} \quad (86d)$$

$$R_0 + R_1 + R_2 \leq I(U; Y_1|W) + I(V; Y_2|W, U) + \min \{I(W; Y_1), I(W; Y_2)\} \quad (86e)$$

over all  $(W, U, V) - X - (Y_1, Y_2)$ , is an outer bound on  $\mathcal{C}$ . By removing the rate bounds in (86c) and (86e) from  $\mathcal{R}_O(0, \infty)$ , one recovers the region  $\mathcal{O}$  from (58). Clearly  $\mathcal{R}_O(0, \infty) \subseteq \mathcal{O}$ , which shows that  $\mathcal{C} \subseteq \mathcal{O}$  and establishes Bound 3.

When there is no common message, one obtains Bound 1, i.e., that  $\hat{\mathcal{C}} \subseteq \hat{\mathcal{O}}$ , by setting  $R_0 = 0$  into Bound 3. This follows by noting that

$$I(V; Y_2|W) + \min \{I(W; Y_1), I(W; Y_2)\} \leq I(W, V; Y_2), \quad (87)$$

and defining  $\tilde{V} = (W, V)$ .

## B. Inner Bounds 2 and 4

Referring to [21, Theorem 1], we have  $\mathcal{R}_I(0, \infty)$  as an inner bound on  $\mathcal{C}$ , where  $\mathcal{R}_I(0, \infty)$  is the closure of the union of rate triples  $(R_0, R_1, R_2) \in \mathbb{R}_+^3$  satisfying:

$$R_1 \leq I(U; Y_1|W) - I(U; V|W) - I(U; Y_2|W, V) \quad (88a)$$

$$R_0 + R_1 \leq I(W, U; Y_1) - I(U; V|W) - I(U; Y_2|W, V) \quad (88b)$$

$$R_0 + R_2 \leq I(W, V; Y_2) \quad (88c)$$

$$R_0 + R_1 + R_2 \leq I(W, U; Y_1) + I(V; Y_2|W) - I(U; V|W) - I(U; Y_2|W, V) \quad (88d)$$

$$R_0 + R_1 + R_2 \leq I(U; Y_1|W) + I(V; Y_2|W) - I(U; V|W) + \min \{I(W; Y_1), I(W; Y_2)\} \quad (88e)$$

over all  $(W, U, V) - X - (Y_1, Y_2)$ . To see that  $\mathcal{I} \subseteq \mathcal{R}_I(0, \infty)$ , note that combining (59a) and (59b) implies (88b). The rate bound in (88c) follows from (59a) and (59c), while (88d) and (88e) both follow by combining (59a)-(59c). As simple consequence of the above is that Bound 2 is an inner bound on the secrecy-capacity region without a common message, which follows by setting  $R_0 = 0$  and  $W = 0$  into Bound 4.

## APPENDIX B

### EXISTENCE OF MAXIMIZING DISTRIBUTION

A key arguments in the proof of Theorem 4 is the continuity of the nested UCE  $T_{\lambda, \eta}^Q(\mathbf{X}) = (\mathfrak{E}t_{\lambda, \eta}^Q)(\mathbf{X})$  in  $P_{\mathbf{X}}$ . This is follows by Proposition 21 from [15], which we reproduced as follows.

**Proposition 10 (Boundedness and Continuity of UCE)** *Consider the space of all Borel probability distributions on  $\mathbb{R}^t$  endowed with the topology induced by weak convergence<sup>4</sup>. Let  $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$  be a sequence of random variable that satisfies the following two properties: (i)  $\exists \kappa > 1$ ,  $B \in \mathbb{R}^{t \times t}$  such that  $\mathbb{E}[\|\mathbf{X}\mathbf{X}^\top\|^\kappa] \leq B$ ,  $\forall n \in \mathbb{N}$  (i.e., the sequence has a uniformly bounded  $\kappa$ -th moment); (ii)  $\mathbf{X}_n \xrightarrow[n \rightarrow \infty]{w} \mathbf{X}^*$ . If  $g : \mathbb{R}^t \rightarrow \mathbb{R}$  is a bounded real-valued function that satisfies  $g(\mathbf{X}_n) \xrightarrow[n \rightarrow \infty]{} g(\mathbf{X}^*)$ , then its UCE  $F = \mathfrak{E}f$  is bounded and satisfies  $G(\mathbf{X}_n) \xrightarrow[n \rightarrow \infty]{} G(\mathbf{X}^*)$ .*

We start by establishing the boundedness of  $t_{\lambda, \eta}^Q(\mathbf{X})$ .

**Lemma 2 (Boundedness of Nested Concave Envelopes)** *For  $\eta > 1$  and  $\lambda_1, \lambda_2 > 0$  there is a  $B_{\lambda, \eta} \in \mathbf{R}$ , such that  $t_{\lambda, \eta}^Q(\mathbf{X}) \leq B_{\lambda, \eta}$ , for all  $P_{\mathbf{X}}$ .*

*Proof:* By Theorem 4, we have that if  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K$ , then

$$\hat{V}_{\lambda, \eta}^Q(K) = t_{\lambda, \eta}^Q(\mathbf{X}^*), \quad (89)$$

where  $\mathbf{X}^* = \mathbf{X}_1^* + \mathbf{X}_2^*$ , and  $\mathbf{X}_1^*$  and  $\mathbf{X}_2^*$  are independent random variables with  $\mathbf{X}_j^* \sim \mathcal{N}(\mathbf{0}, K_j)$ , for  $j = 1, 2$ , such that  $K^* \triangleq K_1 + K_2 \preceq K$ . Furthermore, by the definition of the UCE and the definition of  $\hat{V}_{\lambda, \eta}^Q(K)$ , (89) implies that for every  $\mathbf{X} \sim P_{\mathbf{X}}$  with  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K$ , we have

$$t_{\lambda, \eta}^Q(\mathbf{X}) \leq T_{\lambda, \eta}^Q(\mathbf{X}) \leq \hat{V}_{\lambda, \eta}^Q(K) = t_{\lambda, \eta}^Q(\mathbf{X}^*). \quad (90)$$

Thus,

$$\begin{aligned} \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} t_{\lambda, \eta}^Q(\mathbf{X}) &\leq \lambda_1 I(\mathbf{X}^*; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}^*; \mathbf{Y}_2) + \lambda_1 S_\eta^Q(\mathbf{X}^*) \\ &\stackrel{(a)}{=} \lambda_1 I(\mathbf{X}^*; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}^*; \mathbf{Y}_2) + \lambda_1 s_\eta^Q(\mathbf{X}_1^*) \\ &\stackrel{(b)}{\leq} \lambda_1 I(\mathbf{X}^*; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}^*; \mathbf{Y}_2) + \lambda_1 C_\eta, \end{aligned} \quad (91)$$

<sup>4</sup>A sequence  $\{X_n\}_{n \in \mathbb{N}}$  of real-valued random variables is said to *converge weakly* (or, equivalently, in distribution) to a random variable  $X$  if  $\lim_{n \rightarrow \infty} F_{X_n}(x) = F(x)$ , for every  $x \in \mathbb{R}$  for which  $F_X$  is continuous, where  $F_{X_n}$  and  $F_X$  are the cumulative distribution functions of  $X_n$  and  $X$ , respectively. This notion of convergence is denoted by  $X_n \xrightarrow[n \rightarrow \infty]{w} X$ .

where (a) is by Theorem 4, while (b) follows from an adaptation of [15, Proposition 19] to the function  $s_\eta^Q(\mathbf{X})$  as defined in (17), which implies that for  $\eta > 1$  there exists a  $C_\eta$ , such that  $s_\eta^Q(\mathbf{X}) \leq C_\eta$ , for all  $P_{\mathbf{X}}$  (see Remark 7). By (91), we have

$$\sup_{\mathbf{X}} t_{\lambda, \eta}^Q(\mathbf{X}) \leq \sup_{0 \leq K: \mathbf{X} \sim \mathcal{N}(\mathbf{0}, K)} \lambda_1 I(\mathbf{X}; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2) + \lambda_1 C_\eta. \quad (92)$$

Let  $\Sigma_j = (\mathbf{G}_j^\top \mathbf{G}_j)^{-1}$ ,  $j = 1, 2$ . For  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, K)$ , we write

$$\begin{aligned} & 2\lambda_1 I(\mathbf{X}; \mathbf{Y}_1) - 2(\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2) \\ &= \lambda_1 \log |\mathbf{I} + \mathbf{G}_1 \mathbf{K} \mathbf{G}_1^\top| - (\lambda_1 + \lambda_2) \log |\mathbf{I} + \mathbf{G}_2 \mathbf{K} \mathbf{G}_2^\top| \\ &= \lambda_1 \log |\mathbf{I} + \mathbf{K} \mathbf{G}_1 \mathbf{G}_1^\top| - (\lambda_1 + \lambda_2) \log |\mathbf{I} + \mathbf{K} \mathbf{G}_2 \mathbf{G}_2^\top| \\ &\stackrel{(a)}{=} -\lambda_1 \log |\Sigma_1| + (\lambda_1 + \lambda_2) \log |\Sigma_2| + \lambda_1 (\log |\Sigma_1 + \mathbf{K}| - \lambda \log |\Sigma_2 + K|), \end{aligned} \quad (93)$$

where (a) is by setting  $\lambda = \frac{\lambda_1 + \lambda_2}{\lambda_1} > 1$ . To bound the last two terms, we use the min-max theorem on eigenvalues:

Let  $\mu_i(\mathbf{A})$  be the  $i$ -th smallest eigenvalue of the symmetric matrix  $\mathbf{A} \in \mathbb{R}^{t \times t}$ , we have

$$\mu_i(\mathbf{A}) = \min_{L_i} \max_{\mathbf{0} \neq \mathbf{u} \in L_i} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = \max_{L_{t+1-i}} \min_{\mathbf{0} \neq \mathbf{u} \in L_{t+1-i}} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}, \quad (94)$$

where  $L_i$  is an  $i$ -dimensional subspace of  $\mathbb{R}^t$ . Since the  $t$ -dimensional subspace of  $\mathbb{R}^t$  is unique (that is,  $L_t = \mathbb{R}_t$ ), we obtain

$$\mu_1(\mathbf{A}) = \max_{L_t} \min_{\mathbf{0} \neq \mathbf{u} \in L_t} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = \min_{\mathbf{0} \neq \mathbf{u} \in L_t} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} \quad (95a)$$

$$\mu_t(\mathbf{A}) = \min_{L_t} \max_{\mathbf{0} \neq \mathbf{u} \in L_t} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = \max_{\mathbf{0} \neq \mathbf{u} \in L_t} \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}. \quad (95b)$$

The RHSs of (95) implies that for every non-zero  $\mathbf{u} \in \mathbb{R}^t$  we have  $\mu_1(\mathbf{A}) \leq \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} \leq \mu_t(\mathbf{A})$ . We upper and lower bound the  $i$ -th eigenvalue of  $\mathbf{K} + \Sigma_j$ , for  $j = 1, 2$ , as follows

$$\begin{aligned} \mu_i(\mathbf{K} + \Sigma_j) &= \min_{L_i} \max_{\mathbf{0} \neq \mathbf{u} \in L_i} \left( \frac{\mathbf{u}^\top \mathbf{K} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} + \frac{\mathbf{u}^\top \Sigma_j \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} \right) \\ &\begin{cases} \geq \min_{L_i} \max_{\mathbf{0} \neq \mathbf{u} \in L_i} \left( \frac{\mathbf{u}^\top \mathbf{K} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} + \mu_1(\Sigma_j) \right) = \mu_i(\mathbf{K}) + \mu_1(\Sigma_j), \\ \leq \min_{L_i} \max_{\mathbf{0} \neq \mathbf{u} \in L_i} \left( \frac{\mathbf{u}^\top \mathbf{K} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} + \mu_t(\Sigma_j) \right) = \mu_i(\mathbf{K}) + \mu_t(\Sigma_j). \end{cases} \end{aligned} \quad (96)$$

Hence the eigenvalues of  $\mathbf{K} + \Sigma_j$ ,  $j = 1, 2$ , satisfy

$$\mu_i(\mathbf{K}) + \mu_1(\Sigma_j) \leq \mu_i(\mathbf{K} + \Sigma_j) \leq \mu_i(\mathbf{K}) + \mu_t(\Sigma_j), \quad (97)$$

where  $i \in [1 : t]$ . We now bound the last two terms in (93) as

$$\log |\Sigma_1 + \mathbf{K}| - \lambda \log |\Sigma_2 + K| = \sum_{i=1}^t \log \left( \frac{\mu_i(\mathbf{K} + \Sigma_1)}{\mu_i(\mathbf{K} + \Sigma_2)^\lambda} \right)$$



$$\begin{aligned}
& \stackrel{(a)}{\leq} \sum_{i=1}^t \log \left( \frac{\mu_i(\mathbf{K}) + \mu_t(\Sigma_1)}{(\mu_i(\mathbf{K}) + \mu_1(\Sigma_2))^\lambda} \right) \\
& \leq t \cdot \max_i \log \left( \frac{\mu_i(\mathbf{K}) + \mu_t(\Sigma_1)}{(\mu_i(\mathbf{K}) + \mu_1(\Sigma_2))^\lambda} \right) \\
& \stackrel{(b)}{\leq} t \cdot \log \left( \frac{\mu^* + \mu_t(\Sigma_1)}{(\mu^* + \mu_1(\Sigma_2))^\lambda} \right), \tag{98}
\end{aligned}$$

where (a) follows from by (96), and (b) is by setting  $\mu^* = \max \left\{ 0, \frac{1}{1-\lambda} (\mu_1(\Sigma_2) - \lambda \mu_t(\Sigma_1)) \right\}$  and noting that  $\mu_i(\mathbf{K}) \geq 0$  and that the derivative of the function  $c(x) \triangleq \log(x + \mu_t(\Sigma_1)) - \lambda \log(x + \mu_1(\Sigma_2))$  is zero at  $x = \mu^*$ , negative when  $x > \mu^*$  and positive when  $x < \mu^*$ . By noting that  $\mu_t(\Sigma_1), \mu_1(\Sigma_2) > 0$  (which holds since the positive semidefinite matrices  $\Sigma_1$  and  $\Sigma_2$  are invertible), we conclude that for every  $\mathbf{X} \sim P_{\mathbf{X}}$

$$t_{\lambda, \eta}^Q(\mathbf{X}) \leq -\lambda_1 \log |\Sigma_1| + (\lambda_1 + \lambda_2) \log |\Sigma_2| + \lambda_1 \cdot t \cdot \log \left( \frac{\mu^* + \mu_t(\Sigma_1)}{(\mu^* + \mu_1(\Sigma_2))^{\frac{\lambda_1 + \lambda_2}{\lambda_1}}} \right) \triangleq B_{\lambda, \eta} < \infty. \tag{99}$$

Next, we show that  $t_{\lambda, \eta}^Q(\mathbf{X})$  satisfies the desired continuity property.

**Lemma 3 (Continuity of Nested Concave Envelopes)** *Consider the space of all Borel probability distributions on  $\mathbb{R}^t$  endowed with the topology induced by weak convergence and let  $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$  be a sequence of random variable that satisfies the following two properties: (i)  $\exists \kappa > 1, B \in \mathbb{R}^{t \times t}$  such that  $\mathbb{E}[\|\mathbf{X}\mathbf{X}^\top\|^\kappa] \leq B \forall n$ ; (ii)  $\mathbf{X}_n \xrightarrow[n \rightarrow \infty]{w} \mathbf{X}^*$ . Then  $t_{\lambda, \eta}^Q(\mathbf{X}_n) \xrightarrow[n \rightarrow \infty]{} t_{\lambda, \eta}^Q(\mathbf{X}^*)$ .*

*Proof:* The proof of Lemma 3 follows immediately by an adaptation of [15, Proposition 20] (see Remark 7) and by [15, Theorem 5], which, respectively, imply that  $S_\eta^q(\mathbf{X}_n) \xrightarrow[n \rightarrow \infty]{} S_\eta^q(\mathbf{X}^*)$ , and that  $h(\mathbf{Y}_{j,n}) \xrightarrow[n \rightarrow \infty]{} h(\mathbf{Y}_j^*)$ , for  $j = 1, 2$ . Here  $\mathbf{Y}_{1,n}$  and  $\mathbf{Y}_{2,n}$  are the outputs of the MIMO Gaussian BC with input  $\mathbf{X}_n$ . ■

Based on Lemmas 2 and 3, Proposition 21 from [15] states that  $T_{\lambda, \eta}^Q(\mathbf{X}_n)$  is bounded and that it satisfies  $T_{\lambda, \eta}^Q(\mathbf{X}_n) \xrightarrow[n \rightarrow \infty]{} T_{\lambda, \eta}^Q(\mathbf{X}^*)$ . Having this, the existence of the maximizer of  $\tilde{V}_{\lambda_0, \alpha, \eta}(\mathbf{K})$  is established as follows. Let  $\hat{K} \succeq 0$  and define

$$\begin{aligned}
\tilde{v}_{\lambda_0, \alpha, \eta}(\hat{K}) & \triangleq \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \hat{K}} f_{\lambda_0, \alpha, \eta}(\mathbf{X}) \\
& = \sup_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \hat{K}} (\lambda_2 - \bar{\alpha} \lambda_0) I(\mathbf{X}; \mathbf{Y}_2) - \alpha \lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + T_{\lambda, \eta}^Q(\mathbf{X}). \tag{100}
\end{aligned}$$

Let  $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$  be a sequence of random variables such that  $\mathbb{E}[\mathbf{X}_n \mathbf{X}_n^\top] = \hat{K}$  and  $f_{\lambda_0, \alpha, \eta}(\mathbf{X}_n) \uparrow \tilde{v}_{\lambda_0, \alpha, \eta}(\hat{K})$ , as  $n \rightarrow \infty$ . By [15, Proposition 17] and since  $\mathbb{E}[\mathbf{X}_n \mathbf{X}_n^\top] = \hat{K}$  for every  $n \in \mathbb{N}$ , we have that  $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$  is a tight sequence<sup>5</sup>, and that there exist an  $\mathbf{X}_{\hat{K}}^*$  and a convergent subsequence  $\{\mathbf{X}_{n_m}\}_{m \in \mathbb{N}}$  such that  $\mathbf{X}_{n_m} \xrightarrow[m \rightarrow \infty]{w} \mathbf{X}_{\hat{K}}^*$ . Invoking [15, Proposition 18] once more we have that  $h(\mathbf{Y}_{j, n_m}) \xrightarrow[m \rightarrow \infty]{} h(\mathbf{Y}_{j, \hat{K}}^*)$ , for  $j = 1, 2$ , where  $\mathbf{Y}_{1, n_m}, \mathbf{Y}_{2, n_m}, \mathbf{Y}_{1, \hat{K}}^*$  and

<sup>5</sup>As defined in [15], a sequence of random variables  $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$  taking values in  $\mathbb{R}^t$  is *tight* if for every  $\epsilon > 0$  there exists a compact set  $\mathcal{C}_\epsilon \subset \mathbb{R}^t$ , such that  $\mathbb{P}(\mathbf{X}_n \notin \mathcal{C}_\epsilon) \leq \epsilon, \forall n \in \mathbb{N}$ .

$\mathbf{Y}_{2,\hat{K}}^*$  are the corresponding outputs. Thus,

$$f_{\lambda_0,\alpha,\eta}(\mathbf{X}_{\hat{K}}^*) = \tilde{v}_{\lambda_0,\alpha,\eta}(\hat{K}). \quad (101)$$

By the definition of  $\tilde{V}_{\lambda_0,\alpha,\eta}^Q(K)$ , we write

$$\tilde{V}_{\lambda_0,\alpha,\eta}^Q(K) = \sup_{\substack{(V,\mathbf{X}): \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} f_{\lambda_0,\alpha,\eta}^Q(\mathbf{X}|V) = \sup_{\substack{(V,\mathbf{X}): \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} \sum_v P(v) f_{\lambda_0,\alpha,\eta}^Q(\mathbf{X}|V=v). \quad (102)$$

Since  $\tilde{V}_{\lambda_0,\alpha,\eta}^Q(K)$  is given by a convex combination as above, to obtain the maximizer subject to the covariance constraint it suffices to restrict the ourselves to the family of maximizers  $\mathbf{X}_{\hat{K}}^*$  for  $K \succeq 0$ . Thus,

$$\tilde{V}_{\lambda_0,\alpha,\eta}^Q(K) = \sup_{\substack{\{\alpha_i\}, \{\hat{K}_i\}: \alpha_i \geq 0 \\ \sum_i \alpha_i = 1, \sum_i \alpha_i \hat{K}_i \preceq K}} \sum_i \alpha_i \tilde{v}_{\lambda_0,\alpha,\eta}^Q(\hat{K}_i). \quad (103)$$

It takes  $\frac{t(t+1)}{2}$  constraints to preserve the covariance matrix (due to its symmetry) and one other constraint to preserve  $\sum_i \alpha_i \tilde{v}_{\lambda_0,\alpha,\eta}^Q(\hat{K}_i)$ . Hence, by using the Bunt-Carathedory theorem [23], we can restrict ourselves to convex combinations of at most  $m \triangleq \frac{t(t+1)}{2} + 1$  points, i.e.,

$$\tilde{V}_{\lambda_0,\alpha,\eta}^Q(K) = \sup_{\substack{\{\alpha_i\}, \{\hat{K}_i\}: \alpha_i \geq 0 \\ \sum_{i=1}^m \alpha_i = 1, \sum_{i=1}^m \alpha_i \hat{K}_i \preceq K}} \sum_{i=1}^m \alpha_i \tilde{v}_{\lambda_0,\alpha,\eta}^Q(\hat{K}_i). \quad (104)$$

Consider any sequence of convex combinations  $\left\{ \left\{ \alpha_i^{(n)} \right\}_{i \in [1:m]}, \left\{ \hat{K}_i^{(n)} \right\}_{i \in [1:m]} \right\}_{n \in \mathbb{N}}$  that approaches the supremum as  $n \rightarrow \infty$ . The compactness of the  $m$ -dimensional simplex implies that  $\alpha_i^{(n)} \xrightarrow{n \rightarrow \infty} \alpha_i^*$ , for every  $i \in [1:m]$ . Furthermore, we have the following property of the limiting points  $\alpha_i^*$ ,  $i \in [1:m]$ .

**Lemma 4** For any  $i \in [1:m]$ , if  $\alpha_i^* = 0$  then  $\alpha_i^{(n)} \tilde{v}_{\lambda_0,\alpha,\eta}^Q(\hat{K}_i^{(n)}) \xrightarrow{n \rightarrow \infty} 0$ .

*Proof:* Let  $i \in [1:m]$  be such that  $\alpha_i^* = 0$  and note that for every  $\hat{K} \succeq 0$ , we have

$$\begin{aligned} \tilde{v}_{\lambda_0,\alpha,\eta}^Q(\hat{K}) &\stackrel{(a)}{=} f_{\lambda_0,\alpha,\eta}^Q(\mathbf{X}_{\hat{K}}^*) \\ &\stackrel{(b)}{\leq} \lambda_2 I(\mathbf{X}_{\hat{K}}^*; \mathbf{Y}_2) + \hat{B}_{\lambda_0,\alpha,\eta} \\ &\leq \frac{\lambda_2}{2} \log |I + \mathbf{G}_2 \hat{K} \mathbf{G}_2^\top| + \hat{B}_{\lambda_0,\alpha,\eta}, \end{aligned} \quad (105)$$

where (a) follows from (101) and the non-negativity of mutual information, while (b) is since  $T_{\lambda,\eta}^Q(\mathbf{X})$  is bounded. Let  $\{\alpha_i^{(n_k)}\}_{k \in \mathbb{N}}$  be a subsequence of  $\{\alpha_i^{(n)}\}_{n \in \mathbb{N}}$ , such that  $\alpha_i^{(n_k)} > 0$  for every  $k \in \mathbb{N}$  (if there is no such subsequence, the result of Lemma 4 is immediate). Since  $\alpha_i^{(n_k)} \hat{K}_i^{(n_k)} \preceq K$  and  $\alpha_i^{(n_k)} > 0$ , we obtain

$$\hat{K}_i^{(n_k)} \preceq \frac{1}{\alpha_i^{(n_k)}} K, \quad \forall k \in \mathbb{N}. \quad (106)$$

We thus conclude that

$$\begin{aligned}
\alpha_i^{(n_k)} \tilde{\mathbf{v}}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}_i^{(n_k)}) &\stackrel{(a)}{\leq} \alpha_i^{(n_k)} \left( \frac{\lambda_2}{2} \log |I + \mathbf{G}_2 \mathbf{K}_i^{(n_k)} \mathbf{G}_2^\top| + \hat{B}_{\lambda_0, \alpha, \eta} \right) \\
&\stackrel{(b)}{\leq} \alpha_i^{(n_k)} \left( \frac{\lambda_2}{2} \log \left| I + \mathbf{G}_2 \frac{\mathbf{K}}{\alpha_i^{(n_k)}} \mathbf{G}_2^\top \right| + \hat{B}_{\lambda_0, \alpha, \eta} \right) \\
&\stackrel{(c)}{=} \alpha_i^{(n_k)} \left( \frac{\lambda_2}{2} \sum_{j=1}^t \log \left( 1 + \frac{\mu_j}{\alpha_i^{(n_k)}} \right) + \hat{B}_{\lambda_0, \alpha, \eta} \right) \xrightarrow[n \rightarrow \infty]{0},
\end{aligned} \tag{107}$$

where (a) and (b) follow from (105) and (106), respectively, while (c) follows by denoting the eigenvalues of  $\mathbf{G}_2 \mathbf{K} \mathbf{G}_2^\top$  by  $\{\mu_j\}_{j=1}^t$ .  $\blacksquare$

Based on Lemma 4, we assume that  $\min_{i \in [1:m]} \alpha_i^* \triangleq \alpha^* > 0$ , which implies that  $K_i^{(n)} \preceq \frac{2}{\alpha^*} \mathbf{K}$  uniformly in  $i \in [1 : m]$ , for large enough values of  $n$ . Hence, for each  $i \in [1 : m]$  we can find a convergent subsequence  $\{\mathbf{K}_i^{(n_k)}\}_{k \in \mathbb{N}}$ , such that  $\mathbf{K}_i^{(n_k)} \xrightarrow[k \rightarrow \infty]{} \mathbf{K}_i^*$ . Putting these together, we have

$$\tilde{\mathbf{V}}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = \sum_{i=1}^m \alpha_i^* \tilde{\mathbf{v}}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}_i^*), \tag{108}$$

i.e., one can always find a pair of random variables  $(V^*, \mathbf{X}^*)$  with  $|\mathcal{V}^*| \leq \frac{t(t+1)}{2} + 1$ , such that  $\tilde{\mathbf{V}}_{\lambda_0, \alpha, \eta}^Q(\mathbf{K}) = f_{\lambda_0, \alpha, \eta}^Q(\mathbf{X}^* | V^*)$ .

## APPENDIX C

### PROOF OF PROPOSITION 9

The proof relies on a result from [22, Corollary 2], which we reproduce in the following.

**Proposition 11 (Min-Max Interchange)** *Let  $\Lambda_d$  be the  $d$ -dimensional simplex, i.e.,  $\Lambda_d = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^d \mid \sum_{i=1}^d \lambda_i = 1 \right\}$ . Let  $\mathcal{P}$  be a set of PDFs  $P_U$  over a set  $\mathcal{U}$ . Let  $\{T_i(P_U)\}_{i \in [1:d]}$  be a set of functions such that the set  $\mathcal{A}$ , defined by*

$$\mathcal{A} = \left\{ \mathbf{a} \in \mathbb{R}^d \mid \forall i \in [1 : d], \exists P_U \in \mathcal{P}, a_i \leq T_i(P_U) \right\} \tag{109}$$

*is a convex set. Then*

$$\sup_{P_U \in \mathcal{P}} \min_{\boldsymbol{\lambda} \in \Lambda_d} \sum_{i=1}^d \lambda_i T_i(P_U) = \min_{\boldsymbol{\lambda} \in \Lambda_d} \sup_{P_U \in \mathcal{P}} \sum_{i=1}^d \lambda_i T_i(P_U). \tag{110}$$

Let  $d = 2$  and  $\mathcal{P}$  be the set of PDFs  $P_{W,V,U,\mathbf{X}}$  that satisfy  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{K}$ . Set

$$T_1(P_{W,V,U,\mathbf{X}}) = \lambda_0 I(W; \mathbf{Y}_1) + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W) \tag{111a}$$

$$T_2(P_{W,V,U,\mathbf{X}}) = \lambda_0 I(W; \mathbf{Y}_2) + \lambda_1 \left[ I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \right] + \lambda_2 I(V; \mathbf{Y}_2 | W), \tag{111b}$$

and consider the corresponding set  $\mathcal{A}$  from (109). To show that  $\mathcal{A}$  is convex, let  $(a_1, a_2), (b_1, b_2) \in \mathcal{A}$  and  $P_a, P_b \in \mathcal{P}$  be two PDFs, such that

$$a_i \leq T_i(P_a), \quad b_i \leq T_i(P_b), \quad i = 1, 2. \quad (112)$$

Fix  $\alpha \in [0, 1]$  and consider a PDF  $P$  given by

$$P(w, v, u, \mathbf{x}) = \alpha P_a(w, v, u, \mathbf{x}) + \bar{\alpha} P_b(w, v, u, \mathbf{x}), \quad \forall (w, v, u, \mathbf{x}) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n. \quad (113)$$

Equivalently,  $P$  can be represented by setting  $\tilde{W} = (Q, W)$ , where  $Q \sim \text{Ber}(\alpha)$ , and denoting  $P \triangleq P_{\tilde{W}, V, U, \mathbf{X}} = P_{(Q, W), V, U, \mathbf{X}}$ , for which

$$P_{(Q, W), V, U, \mathbf{X}}((0, w), v, u, \mathbf{x}) = \alpha P_a(w, v, u, \mathbf{x}) \quad (114a)$$

$$P_{(Q, W), V, U, \mathbf{X}}((1, w), v, u, \mathbf{x}) = \bar{\alpha} P_b(w, v, u, \mathbf{x}), \quad (114b)$$

for all  $(w, v, u, \mathbf{x}) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n$ . First note that

$$\mathbb{E}_P[\mathbf{X}\mathbf{X}^T] = \alpha \mathbb{E}_{P_a}[\mathbf{X}\mathbf{X}^T] + \bar{\alpha} \mathbb{E}_{P_b}[\mathbf{X}\mathbf{X}^T] \preceq \mathbf{K}, \quad (115)$$

where  $\mathbb{E}_Q$  denotes that an expectation is taken with respect to a PDF  $Q$ . This implies that  $P \in \mathcal{P}$ .

Next, by evaluating the functions  $T_i$ ,  $i = 1, 2$ , with respect to  $P$ , we have

$$\begin{aligned} T_i(P) &= \lambda_0 I_P(\tilde{W}; \mathbf{Y}_i) + \lambda_1 \left[ I_P(U; \mathbf{Y}_1 | \tilde{W}) - I_P(U; V | \tilde{W}) - I_P(U; \mathbf{Y}_2 | \tilde{W}, V) \right] + \lambda_2 I_P(V; \mathbf{Y}_2 | \tilde{W}) \\ &= \lambda_0 I_P(Q; \mathbf{Y}_i) + \alpha T_i(P_a) + \bar{\alpha} T_i(P_b) \\ &\geq \alpha a_i + \bar{\alpha} b_i, \end{aligned} \quad (116)$$

implying that  $\alpha(a_1, a_2) + \bar{\alpha}(b_1, b_2) \in \mathcal{A}$ , which establishes the convexity of  $\mathcal{A}$ . The proof of Proposition 9 is completed by invoking Proposition 11, while noting that for every tuple of random variables  $(W, V, U, \mathbf{X})$ , with  $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$  and  $\mathbb{E}[\mathbf{X}\mathbf{X}^T] \preceq \mathbf{K}$ ,

$$\min_{\alpha \in [0, 1]} \left\{ \begin{aligned} &\lambda_0 [\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)] \\ &+ \lambda_1 [I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V)] + \lambda_2 I(V; \mathbf{Y}_2 | W) \end{aligned} \right\} \quad (117)$$

is attained by either  $\alpha = 0$  or  $\alpha = 1$ .

## REFERENCES

- [1] Y. Liang, H. V. Poor, and S. Shamai (Shitz). *Foundations and Trends in Communications and Information Theory*, volume 5, chapter Information Theoretic Security, pages 355–580. Now Publishers, MA, USA, 2008.
- [2] Z. Li, W. Trappe, and R. D. Yates. Secret communication via multiantenna transmission. In *Proc. 41st Annu. Conf. Information Sciences and Systems*, Baltimore, Maryland, US, Mar. 2007.
- [3] S. Shafiee, N. Liu, and S. Ulukus. Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel. *IEEE Trans. Inf. Theory*, 55(9):4033–4039, Sep. 2009.
- [4] A. Khisti and G. W. Wornell. Secure transmission with multiple antennas - part I: The MISOME wiretap channel. *IEEE Trans. Inf. Theory*, 56(7):3088–3104, Jul. 2010.
- [5] R. Bustin, R. Liu, H. V. Poor, and S. Shamai. An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel. In *Proc. Int. Symp. Inf. Theory (ISIT-2009)*, pages 2602–2606, Seoul, Korea, Jun.-Jul. 2009.
- [6] T. Liu and S. Shamai. A note on the secrecy capacity of the multiple-antenna wiretap channel. *IEEE Trans. Inf. Theory*, 6(6):2547–2553, Jun. 2009.
- [7] A. Khisti and G. W. Wornell. Secure transmission with multiple antennas - part II: The MIMOME channel. *IEEE Trans. Inf. Theory*, 56(11):5515–5532, Nov. 2010.
- [8] F. Oggier and B. Hassibi. The secrecy capacity of the MIMO wiretap channel. *IEEE Trans. Inf. Theory*, 57(8):4961–4972, Aug. 2011.
- [9] E. Ekrem and S. Ulukus. The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel. *IEEE Trans. Inf. Theory*, 57(4):2083–2114, Apr. 2011.
- [10] H. D. Ly, T. Liu, and Y. Liang. Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages. *IEEE Trans. Inf. Theory*, 56(11):5477–5487, Nov. 2010.
- [11] H.-F. Chong and Y.-C. Liang. Secrecy capacity region of a class of two-user Gaussian MIMO BC with degraded message sets. In *Proc. Int. Symp. Inf. Theory (ISIT-2013)*, pages 2009–2013, Istanbul, Turkey, Jul. 2013.
- [12] R. Liu, T. Liu, H. V. Poor, and S. Shamai. Multiple-input multiple-output Gaussian broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 56(9):4215–4227, Sep. 2010.
- [13] E. Ekrem and S. Ulukus. Capacity region of Gaussian MIMO broadcast channels with common and confidential messages. *IEEE Trans. Inf. Theory*, 58(9):5669–5680, Sep. 2012.
- [14] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Inf. Theory*, 52(9):3936–3964, Sep. 2006.
- [15] Y. Geng and C. Nair. The capacity region of the two-receiver vector Gaussian broadcast channel with private and common messages. *IEEE Trans. Inf. Theory*, 60(4):2087–2094, Apr. 2014.
- [16] P. P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inf. Theory*, 20(2):279–280, Mar. 1974.
- [17] W. Yu and J. Cioffi. Sum capacity of Gaussian vector broadcast channels. *IEEE Trans. Inf. Theory*, 50(9):1875–1892, Sep. 2004.
- [18] A. Khina, Y. Kochman, and A. Khisti. The confidential MIMO broadcast capacity: A simple derivation. In *Proc. Int. Symp. Inf. Theory (ISIT-2015)*, pages 1981–1985, Hong Kong, Jun. 2015.
- [19] Csiszár and Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge Univ. Press, 2nd edition, 2011.
- [20] H. L. Royden. *Real analysis*. Macmillan, 1988.
- [21] Z. Goldfeld, G. Kramer, and H. H. Permuter. Broadcast channels with privacy leakage constraints. *Submitted for publication to IEEE Trans. Inf. Theory*, 2015.
- [22] Y. Geng, A. Gohari, C. Nair, and Y. Yu. On Marton's inner bound and its optimality for classes of product broadcast channels. *IEEE Trans. Inf. Theory*, 60(1):22–41, 2014.
- [23] L. Bunt. *Bijdrage tot de theorie der convexe puntverzamelingen*. PhD thesis, Univ. Groningne, Amsterdam, 1934.